

STIC-Biotech/ChemLib

80355

From: Yu, Misook
Sent: Monday, November 18, 2002 8:26 AM
To: STIC-Biotech/ChemLib
Subject: RE: 09/191,497

09 919 497

Please disregard this request. I give you a wrong serial no. Thank you.

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

-----Original Message-----

From: STIC-Biotech/ChemLib
Sent: Monday, November 18, 2002 8:03 AM
To: Yu, Misook
Subject: RE: 09/191,497

This Application Number does not have a valid CRF - please provide the necessary correction or another application number.

MAUDE

-----Original Message-----

From: Yu, Misook
Sent: Sunday, November 17, 2002 1:30 PM
To: STIC-Biotech/ChemLib
Subject: 09/191,497

Please search SEQ ID NO:14 and 32

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 11/18
Date Completed: 11/27
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

**PALM INTRANET**Day : Monday
Date: 11/18/2002
Time: 07:19:25

Biotech Query for 09/191497

Title: **WEB-BASED ENGINEERING DESIGN AGENTS**Inventor: **RADCLIFFE, CLARK J.**Location: **9200/FILE REPOSITORY (FRANCONIA)**Location Date: **09/20/2001**Group Art Unit: **2177**Status: **150/PATENTED CASE**

Num	Date	Code	Contents Description
NO BIOTECH DATA			

Search for Biotech Info: Application# PCT / /

To go back use Back button on your browser toolbar.

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)*auto 8:03*

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

80355

From: Yu, Misook
Sent: Sunday, November 17, 2002 1:30 PM
To: STIC-Biotech/ChemLib
Subject: 09/191,497

mej

09 919 497

Please search SEQ ID NO:14 and 32

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

TYPE OF SEARCH:

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:44:35 ; Search time 6670.76 Seconds
(without alignments)
16827.091 Million cell updates/sec

Title: US-09-919-497-14

Perfect score: 3857

Sequence: 1 ggaagagcttcacatgagctc.....agaatacaacctaagctca 3857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_bt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3857	100.0	3857	9 HUMERCC4C1	L77890 Homo sapien
2	2866	74.3	2881	9 HSU64315	U64315 Human DNA r
3	2718	70.5	2718	6 AR183109	AR183109 Sequence
4	2718	70.5	2718	6 AR203246	AR203246 Sequence
5	2718	70.5	2718	6 AX277006	AX277006 Sequence
6	1932.2	50.1	2858	10 AB017635	AB017635 Citicellu
7	1886	48.9	3445	10 AF189285	AF189285 Mus muscu
8	1872	48.5	3235	6 BC026792	BC026792 Mus muscu
9	1829.4	47.4	35641	6 AX332256	AX332256 Sequence
10	1829.4	47.4	35641	6 AX409578	AX409578 Sequence
11	1829.4	47.4	35641	9 HUMERCC4G	L76568 Homo sapien
12	1829.4	47.4	119172	9 AC010401	AC010401 Homo sapi
13	1826.2	47.3	149849	9 AC009173	AC009173 Homo sapi
14	1233.4	32.0	29755	9 AR491814	AR491814 Homo sapi
15	1101	28.5	1101	6 AR183108	AR183108 Sequence
16	1101	28.5	1101	6 AR203245	AR203245 Sequence
17	1101	28.5	1101	6 AX277004	AX277004 Sequence
18	803.4	20.8	846	11 G67515	G67515 csnpercc4-P
19	602.4	15.6	718	11 G67516	G67516 csnpercc4-P
20	464.6	12.0	167804	2 AC099214	AC099214 Rattus no
21	459.8	11.9	128026	10 AC004155	AC004155 Mus muscu
22	458.2	11.6	167329	2 AC096307	AC096307 Rattus no
23	448.4	11.6	182509	2 AC117889	AC117889 Rattus no
24	220.4	5.7	312	11 G67513	G67513 csnpercc4-P
25	220.4	5.7	623	9 HSA334998	AJ334998 Homo sapi
26	220.4	5.7	689	9 HSA324979	AJ324979 Homo sapi
27	220.4	5.7	700	9 HSA335478	AJ335478 Homo sapi
28	199.4	5.2	306	11 G67514	G67514 csnpercc4-P
29	171.4	4.4	2064	8 AF191494	AF191494 Arabidops
30	171.4	4.4	3151	8 AF089003	AF089003 Arabidops
31	171.4	4.4	3173	8 AF160500	AF160500 Arabidops
32	159	4.1	160986	3 AC108481	AC108481 Drosophill
33	159	4.1	294218	3 AE003430	AE003430 Drosophill
34	157.4	4.1	3502	3 DM027181	U27181 Drosophill
35	157.4	4.1	5123	2 AC014388	AC014388 Drosophill
36	154.6	4.0	3857	9 HUMERCC4G	L77890 Homo sapien
37	154.6	4.0	29755	9 AF491814	AF491814 Homo sapi
38	154.6	4.0	35641	6 AX332256	AX332256 Sequence
39	154.6	4.0	35641	6 AX409578	AX409578 Sequence
40	154.6	4.0	35641	9 HUMERCC4G	L76568 Homo sapien
41	154.6	4.0	119172	9 AC010401	AC010401 Homo sapi
42	154.6	4.0	149849	9 AC009173	AC009173 Homo sapi
43	151.4	3.9	182509	2 AC117889	AC117889 Rattus no
44	150.8	3.9	4048	8 AF277377	AF277377 Arabidops
45	148.8	3.9	83698	8 AB010072	AB010072 Arabidops

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
HUMERCC4C1	HUMERCC4C1	3857 bp DNA	L77890	L77890	LINEAR	GI:1905923	Homo sapiens adult DNA.	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HUMERCC4G	HUMERCC4G	3857 bp DNA	L77890	L77890	LINEAR	GI:1905923	Homo sapiens adult DNA.	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HSU64315	HSU64315	Human DNA r	U64315	U64315	Human DNA r	U64315	Human DNA r	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AR183109	AR183109	Sequence	AR183109	AR183109	Sequence	AR183109	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AR203246	AR203246	Sequence	AR203246	AR203246	Sequence	AR203246	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX277006	AX277006	Sequence	AX277006	AX277006	Sequence	AX277006	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AB017635	AB017635	Citicellu	AB017635	AB017635	Citicellu	AB017635	Citicellu	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF189285	AF189285	Mus muscu	AF189285	AF189285	Mus muscu	AF189285	Mus muscu	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
BC026792	BC026792	Mus muscu	BC026792	BC026792	Mus muscu	BC026792	Mus muscu	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX332256	AX332256	Sequence	AX332256	AX332256	Sequence	AX332256	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX409578	AX409578	Sequence	AX409578	AX409578	Sequence	AX409578	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HUMERCC4G	HUMERCC4G	L76568 Homo sapien	L76568	L76568	Homo sapien	L76568	Homo sapien	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC010401	AC010401	Homo sapi	AC010401	AC010401	Homo sapi	AC010401	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC009173	AC009173	Homo sapi	AC009173	AC009173	Homo sapi	AC009173	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AR491814	AR491814	Homo sapi	AR491814	AR491814	Homo sapi	AR491814	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AR183108	AR183108	Sequence	AR183108	AR183108	Sequence	AR183108	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AR203245	AR203245	Sequence	AR203245	AR203245	Sequence	AR203245	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX277004	AX277004	Sequence	AX277004	AX277004	Sequence	AX277004	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
G67515	G67515	csnpercc4-P	G67515	G67515	csnpercc4-P	G67515	csnpercc4-P	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
G67516	G67516	csnpercc4-P	G67516	G67516	csnpercc4-P	G67516	csnpercc4-P	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC099214	AC099214	Rattus no	AC099214	AC099214	Rattus no	AC099214	Rattus no	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC004155	AC004155	Mus muscu	AC004155	AC004155	Mus muscu	AC004155	Mus muscu	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC096307	AC096307	Rattus no	AC096307	AC096307	Rattus no	AC096307	Rattus no	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC117889	AC117889	Rattus no	AC117889	AC117889	Rattus no	AC117889	Rattus no	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
G67513	G67513	csnpercc4-P	G67513	G67513	csnpercc4-P	G67513	csnpercc4-P	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HSA334998	HSA334998	Homo sapi	AJ334998	AJ334998	Homo sapi	AJ334998	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HSA324979	HSA324979	Homo sapi	AJ324979	AJ324979	Homo sapi	AJ324979	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
HSA335478	HSA335478	Homo sapi	AJ335478	AJ335478	Homo sapi	AJ335478	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
G67514	G67514	csnpercc4-P	G67514	G67514	csnpercc4-P	G67514	csnpercc4-P	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF191494	AF191494	Arabidops	AF191494	AF191494	Arabidops	AF191494	Arabidops	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF089003	AF089003	Arabidops	AF089003	AF089003	Arabidops	AF089003	Arabidops	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF160500	AF160500	Arabidops	AF160500	AF160500	Arabidops	AF160500	Arabidops	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC108481	AC108481	Drosophill	AC108481	AC108481	Drosophill	AC108481	Drosophill	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AE003430	AE003430	Drosophill	AE003430	AE003430	Drosophill	AE003430	Drosophill	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
U27181	U27181	Drosophill	U27181	U27181	Drosophill	U27181	Drosophill	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC014388	AC014388	Drosophill	AC014388	AC014388	Drosophill	AC014388	Drosophill	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
L77890	L77890	Homo sapien	L77890	L77890	Homo sapien	L77890	Homo sapien	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF491814	AF491814	Homo sapi	AF491814	AF491814	Homo sapi	AF491814	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX332256	AX332256	Sequence	AX332256	AX332256	Sequence	AX332256	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AX409578	AX409578	Sequence	AX409578	AX409578	Sequence	AX409578	Sequence	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
L76568	L76568	Homo sapien	L76568	L76568	Homo sapien	L76568	Homo sapien	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC010401	AC010401	Homo sapi	AC010401	AC010401	Homo sapi	AC010401	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC009173	AC009173	Homo sapi	AC009173	AC009173	Homo sapi	AC009173	Homo sapi	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AC117889	AC117889	Rattus no	AC117889	AC117889	Rattus no	AC117889	Rattus no	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AF277377	AF277377	Arabidops	AF277377	AF277377	Arabidops	AF277377	Arabidops	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	
AB010072	AB010072	Arabidops	AB010072	AB010072	Arabidops	AB010072	Arabidops	Brookman,K.W., Lamerdin,J.E., Thelen,M.P., Hwang,M., Reardon,J.T., Sancar,A., Zhou,Z.Q., Walter,C.A., Parls,C.N. and Thompson,L.H.	

TITLE ERCC4 (XPF) encodes a human nucleotide excision repair protein with
JOURNAL eukaryotic recombination homologs
MOL. CELL. BIOL. 16 (11), 6553-6562 (1996)
MEDLINE 97042484
PUBMED 8887684
REFERENCE 2 (bases 1 to 3857)
AUTHORS Lamerdin, J. E.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1996) J. E. Lamerdin, Human Genome Center,
Lawrence Livermore National Laboratory, 7000 East Ave, Livermore,
CA, USA, 94551 jamee@cgcl.llnl.gov ow@cornak.llnl.gov
COMMENT GSDB:S:109760
FEATURES
source
1. 3857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="16p13.13"
/cell_line="Hela"
/cell_type="fibroblast"
/dev_stage="adult"
13. 2763
/gene="ERCC4"
13. 2763
/note="ERCC4"
/note="putative"
/product="DNA repair protein"
/protein_id="AA050174.1"
/db_xref="GI:1905924"
/db_xref="GDB:119113"
/translation="MESGPARIAMAPLLEYEROLVELLDLTDLVLCARGLAGDRL
LHFLQHCHPACLVLYLNTOPAEVEYFINOKIEVEHLPRVTEINETSNSRYEYV
OGGVFEATSRILVDFLDRIISDLITGLVYRAHRIIESCOEAFILRLFPKRNKGF
IKAFIDNAVAFDTGCHVERVMNLEFRKLYLMPRHVAVNSFLEBQKREVEYIHSM
TPMLAIOTAILDILNACLKELCHNPSEVEDLSLEMTLGPDKTIRHYDPLMHO
LGAKTSIVODIKILRTLOYLSOYDCYFNLBESLREKAFSGNSWFLDSTST
METINARARYNHPDAMSKREKISEKMEIKESBEIKELVLSNPKWEKLVLEKTE
AENKESALGPGQVYLICASDRTCSQLDYITLGEAFLLRLYRTEFKDSKAEEVM
MKFEREDSKRIKSHKRPDKONKERRATKERTLKRRKLLTOMVGKPELEBEG
DVEEGYRREIVSSPESCPKEIKHEEFDVNLSSDAAGLKEPLTIHPLRGSDPYAL
TVLHEVERRYVLYDAELTFVROLEIYASRPGRLARYFPLIGGSTTEORYLALR
KEKAEKELIRKASMYPEREGRDETLDLVRCASADVSTDPKAGSONGNOQ
STVDMREBSRSLPSLIRRGDIEPYLEVGDYILTPMKCYERKISDLISLNGK
LISQISMSTRYKRPVLLIEFDPSPKPSLTSGALFOELISSNDISKLLTLHLHPRL
RLMCPSPHATLELEFEELKQSPQDAATAIATDSETLPESEKYNPQDFLLKMP
GVNANKRSLMHVKNIAELALSDDELTSILGNANAKQLVDYFHTSPAEVYSGKG
KK"
polyA_site
3812
/note="putative"
3845
/note="putative"
polyA_site
3845
/note="putative"
BASE COUNT 1134 a 851 c 850 g 1022 t
ORIGIN
Query Match 100.0% Score 3857: DB 9: Length 3857:
Best Local Similarity 100.0% Pred. No. 0:
Matches 3857: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 GGAAGAGCTTCATGAGTCAGGCGGCTCGACGATGCCATGGCGCGCTGCTG 60
DB 1 GGAAGAGCTTCATGAGTCAGGCGGCTCGACGATGCCATGGCGCGCTGCTG 60
QY 61 GACTACGAGGACAGCTGGTGTGTAAGTGTGACACTGACGCGGCTAGTAGTGGCGC 120
DB 61 GACTACGAGGACAGCTGGTGTGTAAGTGTGACACTGACGCGGCTAGTAGTGGCGC 120
QY 121 CGGAGGCTGGGCGGCGGCTCTACACTTTCTCCAGCTGACGCGCAGCC 180
DB 121 CGGAGGCTGGGCGGCGGCTCTCTACACTTTCTCCAGCTGACGCGCAGCC 180
QY 181 TGGCTGTGCTGCTCAACAGCAGCGCGGAGAGAGATTTATCATCACTG 240
DB 181 TGGCTGTGCTGCTCAACAGCAGCGCGGAGAGAGATTTATCATCACTG 240

QY 241 AAGTAGAGAGGTGAACACCTCCCTGCGGTACAAATGAATCACAGCAAGT 300
DB 241 AAGTAGAGAGGTGAACACCTCCCTGCGGTACAAATGAATCACAGCAAGT 300
QY 301 CGGTATGAGTTTACACACAAGGTGTATATTTGCGACAAGTAGATCTGTGTT 360
DB 301 CGGTATGAGTTTACACACAAGGTGTATATTTGCGACAAGTAGATCTGTGTT 360
QY 361 GACTTCTTGACTGATAGAAATACCTTCAGATTTAATTAAGTGCATCTTGGTATAGACC 420
DB 361 GACTTCTTGACTGATAGAAATACCTTCAGATTTAATTAAGTGCATCTTGGTATAGACC 420
QY 421 CACAGATTAATCGAGTGTTCAGAAAGCATATCTTGGCGCTTTCGCCGAAAAAC 480
DB 421 CACAGATTAATCGAGTGTTCAGAAAGCATATCTTGGCGCTTTCGCCGAAAAAC 480
QY 481 AAGCTGTGTTTATTAAGCTTTCACAGACAATGCTTGTCCCTTGTATCTGTTTGT 540
DB 481 AAGCTGTGTTTATTAAGCTTTCACAGACAATGCTTGTCCCTTGTATCTGTTTGT 540
QY 541 CATGTGAAAGAGTGAAGAAATCTTTTGTGAGGAACTGTATCTGTGCCAAGTTC 600
DB 541 CATGTGAAAGAGTGAAGAAATCTTTTGTGAGGAACTGTATCTGTGCCAAGTTC 600
QY 601 CATGTAGCACTAAACATCTTTTGAACACAGCAAACTGAGTTCAGAAATCATGTT 660
DB 601 CATGTAGCACTAAACATCTTTTGAACACAGCAAACTGAGTTCAGAAATCATGTT 660
QY 661 TGTATGACACCTTCCATCATGCTGTCTATACAGCTGTATATCTGACATTTTAAATGCATGT 720
DB 661 TGTATGACACCTTCCATCATGCTGTCTATACAGCTGTATATCTGACATTTTAAATGCATGT 720
QY 721 CTAAAGAACTAAAGCCATTAACCATGCTTGAAGTGAAGATTAATTAAGAAAT 780
DB 721 CTAAAGAACTAAAGCCATTAACCATGCTTGAAGTGAAGATTAATTAAGAAAT 780
QY 781 GCTATTGGAAGACCTTTTGAACAAGACAATCGCATTTATCTGATCTTGTGGACACAG 840
DB 781 GCTATTGGAAGACCTTTTGAACAAGACAATCGCATTTATCTGATCTTGTGGACACAG 840
QY 841 CTGTGAGCCAAAGCTTAATCTTATGTTCAAGATTTGAAGATTTAGCACTTTGCTGAC 900
DB 841 CTGTGAGCCAAAGCTTAATCTTATGTTCAAGATTTGAAGATTTAGCACTTTGCTGAC 900
QY 901 TATCTCTCTAGATGATGTTGTGCACATTTCTTAATCTTGAATCTTGAAGCAAGC 960
DB 901 TATCTCTCTAGATGATGTTGTGCACATTTCTTAATCTTGAATCTTGAAGCAAGC 960
QY 961 GAAAAAGCTTTTGTGCAAGATTCAGTTGGCTTTTCTTGACTCCAGCACCTGATGTTT 1020
DB 961 GAAAAAGCTTTTGTGCAAGATTCAGTTGGCTTTTCTTGACTCCAGCACCTGATGTTT 1020
QY 1021 ATTAATCTGACAGCAAGGTTTATCATCTTCCAGATGCCAAATGATTAAGAAAAA 1080
DB 1021 ATTAATCTGACAGCAAGGTTTATCATCTTCCAGATGCCAAATGATTAAGAAAAA 1080
QY 1081 ATATCTGAAAAAATGAAAAATTAAGAGGAGAAAAAAGCAACTGCTCTAGAA 1140
DB 1081 ATATCTGAAAAAATGAAAAATTAAGAGGAGAAAAAAGCAACTGCTCTAGAA 1140
QY 1141 AGCAACCCAAAGTGGGAGGCACTGACTGAAGTATTAAGAAATTTGAGCAGAAAAATAG 1200
DB 1141 AGCAACCCAAAGTGGGAGGCACTGACTGAAGTATTAAGAAATTTGAGCAGAAAAATAG 1200
QY 1201 GAGAGTGAAGCTCTGTTGTTGTCAGGCAAGTCTGATTTGTGCAAGTGTGACCCAGCA 1260
DB 1201 GAGAGTGAAGCTCTGTTGTTGTCAGGCAAGTCTGATTTGTGCAAGTGTGACCCAGCA 1260
QY 1261 TGTTCACAGCTGAGAGACTATATCACTCTTGGAGCGGAGGCTTCTTATTTAGAGCTCTAC 1320
DB 1261 TGTTCACAGCTGAGAGACTATATCACTCTTGGAGCGGAGGCTTCTTATTTAGAGCTCTAC 1320

1321 AGAAAACTTTGAGAGATGACAAAGCTGAGAGTGTGATGTAATTTAGAGAGAA 1380
1322 AGAAAACTTTGAGAGATGAGCAAGCTGAGAGTGTGATGTAATTTAGAGAGAA 1380
1381 GACAGTTCAGAGAAATTTAGAAATCTCAGAAAGACTTAAAGACCCCAAAAGAA 1440
1382 GACAGTTCAGAGAAATTTAGAAATCTCAGAAAGACTTAAAGACCCCAAAAGAA 1440
1441 GGGGCTTACCAAGAAAGAAACCTCAAAAGAAAGAAAGAGTGTGACTTAATCTCA 1500
1442 GGGGCTTACCAAGAAAGAAAGAAAGAAAGAAAGAAAGTGTGACTTAATCTCA 1500
1501 ATGTAGAGAAAGCTGAGAAAGCTGAGAGAGAGATGTGAGAGAGATATGCTGCA 1560
1502 ATGTAGAGAAAGCTGAGAAAGCTGAGAGAGAGATGTGAGAGAGATATGCTGCA 1560
1561 GAAATAG 1620
1562 GAAATAG 1620
1621 AATTTGATGAG 1680
1622 AATTTGATGAG 1680
1681 CTTCTGAG 1740
1682 CTTCTGAG 1740
1741 TACGAG 1800
1742 TACGAG 1800
1801 AGTAG 1860
1802 AGTAG 1860
1861 CAAGCCTATCTACGCTTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
1862 CAAGCCTATCTACGCTTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
1921 AAG 1980
1922 AAG 1980
1981 GTAG 2040
1982 GTAG 2040
2041 CAG 2100
2042 CAG 2100
2101 TCTCTGAG 2160
2102 TCTCTGAG 2160
2161 ATCTCTGAG 2220
2162 ATCTCTGAG 2220
2221 AATAG 2280
2222 AATAG 2280
2281 CTTCTGAG 2340
2282 CTTCTGAG 2340
2341 CAG 2400
2342 CAG 2400
2401 AGACTAG 2460

2401 AGACTAG 2460
2461 AAACAAAG 2520
2462 AAACAAAG 2520
2521 ACCCTTCCGAG 2580
2522 ACCCTTCCGAG 2580
2581 GGGGAG 2640
2582 GGGGAG 2640
2641 GCAG 2700
2642 GCAG 2700
2701 CTTTATGAG 2760
2702 CTTTATGAG 2760
2761 TGAG 2820
2762 TGAG 2820
2821 CATCATAG 2880
2822 CATCATAG 2880
2881 ATTGAG 2940
2882 ATTGAG 2940
2941 AACTGAG 3000
2942 AACTGAG 3000
3001 CTTTAAATGAG 3060
3002 CTTTAAATGAG 3060
3061 AACTTAAATGAG 3120
3062 AACTTAAATGAG 3120
3121 CGGTGAG 3180
3122 CGGTGAG 3180
3181 TCTAG 3240
3182 TCTAG 3240
3241 AATGAG 3300
3242 AATGAG 3300
3301 TGTGAG 3360
3302 TGTGAG 3360
3361 CCATTCGAG 3420
3362 CCATTCGAG 3420
3421 TCTAAG 3480
3422 TCTAAG 3480
3481 GTTACAG 3540

Db 3481 GTTGACTAGCTCTTCTAATGACCCCTAACAATATAGCATATACATATATTCATTTGTTCCA 3540
Qy 3541 AATTAGTATTTTAAAGCAAAATGAAATTCCTGTTGCAAAAGTTAATGATGAAGAGCT 3600
Db 3541 AATTAGTATTTTAAAGCAAAATGAAATTCCTGTTGCAAAAGTTAATGATGAAGAGCT 3600
Qy 3601 CTAGAAATCTCAATTTTGGACATATTCAGTCTCCTAATATCAGAGATCCCTAAGTCCA 3660
Db 3601 CTAGAAATCTCAATTTTGGACATATTCAGTCTCCTAATATCAGAGATCCCTAAGTCCA 3660
Qy 3661 GCTGCTAGTACAGAGATTTTTCAGACTTCTCCTGTTTTCAGAGCTTTATTCCTAAGAC 3720
Db 3661 GCTGCTAGTACAGAGATTTTTCAGACTTCTCCTGTTTTCAGAGCTTTATTCCTAAGAC 3720
Qy 3721 ACCAGATATATTCCTCTGAAATACACCTAATGCGAGTGGCGGAGTGGCAGCT 3780
Db 3721 ACCAGATATATTCCTCTGAAATACACCTAATGCGAGTGGCGGAGTGGCAGCT 3780
Qy 3781 GCACCCCTGCTGCGGACAGAGTGAAGTCTGCTCTATTCACAAAAGAAAAGAAAAGA 3840
Db 3781 GCACCCCTGCTGCGGACAGAGTGAAGTCTGCTCTATTCACAAAAGAAAAGAAAAGA 3840
Qy 3841 AATACAACTAAGCTCA 3857
Db 3841 AATACAACTAAGCTCA 3857

RESULT 2
HSU64315 2881 bp mRNA linear PRI 06-SEP-1996
LOCUS Human DNA repair endonuclease subunit (XPF) mRNA, complete cds.
DEFINITION U64315
ACCESSION U64315.1 GI:1524410
VERSION ERCC4; ERCC11; nucleotide excision repair; xeroderma pigmentosum.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2881)
Sijbers,A.M., de Laat,W.L., Ariza,R.R., Biggelaar,M., Wei,Y.,
Moggs,J.G., Carter,K.C., Shell,B.K., Evans,E., de Jong,M.C.,
Rademakers,S., de Kooij,J., Jaspers,N.G., Hoeijmakers,J.H. and
Wood,R.D.
Xeroderma pigmentosum group F caused by a defect in a
structure-specific DNA repair endonuclease
Cell (1996) in press
2 (bases 1 to 2881)
Wood,R.D.
Direct Submission
Submitted (18-JUL-1996) Biochemistry of Inherited Syndromes Lab.,
Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche
Lane, South Mimms, Herts., EN6 3LD, United Kingdom

FEATURES
source
1..2881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.1-13.2"
1..2881
/gene="XPF"
16..2733
/gene="XPF"
/codon_start=1
/product="DNA repair endonuclease subunit"
/protein_id="AA07689.1"
/db_xref="GI:1524411"
/translation="MAPLEDEROLVELLDLDGLVVCARGLGAADRLYLHFLQNHCP
ACLVLVNTPAEVEEYINOLKIEGVEHLPRVTNTITSREYVTOGVIATSRIL
LYVDELDTLPDLITGLIVRAHRTIESGOEAFILIRROKKBGFIKAFIDNAVAF
DVGFCHEVRMRLPYRKLYLMPRHVAVANSLEQIKRPEVEYELHVSMTPLMLAIGAI
LDLNLACLEKLNHPDSLEDAIGKRPDKTIRHLDELHMLQGLKTKSLVD
LKIIRTLQIYLSQYDCVTFNLLESIRATEKAGONGSGLFDSSTSMFINARAVYH
LPDAKMSKKEIKSEKMEIKGEETKKEIVLESNPKWELTEVLEKTEAEKSEALG

PGOVLICASDRTCSQLRDYITLGAENFLRLRYKTEKDSKAEEVMMKREKEDSKR
IRKSHKRPKDKPONKERRASTKERTLKKKKRLJLQMGKPELEEGVEGIRREIS
SSPESCPREKIEHEFDVNLSSDAAGIKERLJLHPLGSDPALRVLEVEPRY
VVIDAELTFVROLEIYRASRPGKRLRYFLYGGSPTEORLYALREKAEFLKIR
EKASNVPEBERGDEFTLDLVRGTSADVSTDPYKAGGONGTOOSIYVDMREFS
ELPSLIRHRDIDIEVYTLVEGDYIITPEKCYERKSISLISLNNGRYSOCISMR
YKRPVLLIEFDPSPKRFSLTSRGALEQELSSNDJISKLTLHLHPRLILMCPSPHAT
ALEFEELKOSKROPDAATALAITADSEFLPSEKRYKCPQDFLLKMPGVNAKNRSLM
HHVKNIAELIALISOBELTSLIGNANAKQLVDPIHLSFAEVVSKGKKK"
2090
variation
/gene="XPF"
/replace="G"
2881..2290
variation
/gene="XPF"
/replace="T"
2377
variation
/gene="XPF"
/replace="T"
2487
variation
/gene="XPF"
/replace="T"
BASE COUNT 861 a 627 c 667 g 726 t
ORIGIN
Query Match 74.3%; Score 2866; DB 9; Length 2881;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 31 GCCTGACGGATTGCCATGGCGCCGCTGCTGAGTACGAGACACCTGGTGTGTAACATG 90
Db 1 GCTGACGGATTGCCATGGCGCCGCTGCTGAGTACGAGACACCTGGTGTGTAACATG 60
Qy 91 CTCGACACTGACGCGCTAGTAGTGTGCGCCGCGGCGCTGCGGACCGGCTCTCTAC 150
Db 61 CTCGACACTGACGCGCTAGTAGTGTGCGCCGCGGCGCTGCGGACCGGCTCTCTAC 120
Qy 151 CACTTCTCGAGCTGCACAGTCCACCCAGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 210
Db 121 CACTTCTCGAGCTGCACAGTCCACCCAGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 211 GCCGAGGAGAGATTATTCATCAATCAGCTGAATAGTAAGATGTAACACCTCCCTCGC 270
Db 181 GCCGAGGAGAGATTATTCATCAATCAGCTGAATAGTAAGATGTAACACCTCCCTCGC 240
Qy 271 CGGTAAACAATGAATACACAGCAACAGCTGCTATGAACTTACACAGAGTGTGT 330
Db 241 CGGTAAACAATGAATACACAGCAACAGCTGCTATGAACTTACACAGAGTGTGT 300
Qy 331 ATATTTCGACAGTAGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390
Db 301 ATATTTCGACAGTAGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 391 TTATATTGAGCATCTGGTGTATAGAGCCACAGATTAATGAGTCTTGTCAAGAAGCA 450
Db 361 TTATATTGAGCATCTGGTGTATAGAGCCACAGATTAATGAGTCTTGTCAAGAAGCA 420
Qy 451 TTGATCTTGGCGCTCTTTCGCCAGAAAACAAACGTTTATTAAGCTTTTCACAGAC 510
Db 421 TTGATCTTGGCGCTCTTTCGCCAGAAAACAAACGTTTATTAAGCTTTTCACAGAC 480
Qy 511 AATGCTGTGCTTGTATGATGCTTTTGTGATGTGGAAGAAGTGAAGAAATCTTTT 570
Db 481 AATGCTGTGCTTGTATGATGCTTTTGTGATGTGGAAGAAGTGAAGAAATCTTTT 540
Qy 571 GTAGGAACTGTATCTGTGGCCAGGTTCCATGTATGACGTAATCTATTTTGAAGAC 630
Db 541 GTAGGAACTGTATCTGTGGCCAGGTTCCATGTATGACGTAATCTATTTTGAAGAC 600
Qy 631 CACAAACCTGAAGTTGAGAAATCATGTTTCTATGACACCTACCATGCTTGTATACAG 690
Db 601 CACAAACCTGAAGTTGAGAAATCATGTTTCTATGACACCTACCATGCTTGTATACAG 660
Qy 691 ACTGCTATGACATTTAAATGCATGTCTAAAGAACTAAATGCCATACCCATCG 750

Db 661 ACTGCTATACGACATTTTAAATGCAATGCTAAAGAGACTAAATATGCATATACCCATCG 720
Qy 751 CTTGAAGTGAAGATTTATCTTTAGAAAATGCTATTTGAAAACTTTTGACAAAGCAATC 810
Db 721 CTTGAAGTGAAGATTTATCTTTAGAAAATGCTATTTGAAAACTTTTGACAAAGCAATC 780
Qy 811 CGCATATATCTGATCTTTTGTGACACAGCTTTGGAGCAGACTTAAATCTTATGTTGAG 870
Db 781 CGCATATATCTGATCTTTTGTGACACAGCTTTGGAGCAGACTTAAATCTTATGTTGAG 840
Qy 871 GATTGAAGATTTAGCAACTTTGTCAGATCTCTCTGATGATGATGTCATATTT 930
Db 841 GATTGAAGATTTAGCAACTTTGTCAGATCTCTCTGATGATGATGTCATATTT 900
Qy 931 CTTAATCTCTGATCTCTGAGCAACGAGAAAAAGCTTTTGTGATGATGATGATGATG 990
Db 901 CTTAATCTCTGATCTCTGAGCAACGAGAAAAAGCTTTTGTGATGATGATGATGATG 960
Qy 991 CTGTTCTGATCTGACGACCTGATGTTTAAATGCTGAGCAAGGCTTATCATCTT 1050
Db 961 CTGTTCTGATCTGACGACCTGATGTTTAAATGCTGAGCAAGGCTTATCATCTT 1020
Qy 1051 CCAGATGCCAAATGAGTAAAAAGAAAAATATCTGAAAAATGGAATTTAAAGAAAGG 1110
Db 1021 CCAGATGCCAAATGAGTAAAAAGAAAAATATCTGAAAAATGGAATTTAAAGAAAGG 1080
Qy 1111 GAAGAAACAAAAAGAACTGCTCTAGAAAAGCAACCAAGTGGAGGCACTGACTAA 1170
Db 1081 GAAGAAACAAAAAGAACTGCTCTAGAAAAGCAACCAAGTGGAGGCACTGACTAA 1140
Qy 1171 GTATTAAAGAAATGAGCGAAAAATTAAGAGAGTGAAGCTCTTGAGTGCAGCTCAA 1230
Db 1141 GTATTAAAGAAATGAGCGAAAAATTAAGAGAGTGAAGCTCTTGAGTGCAGCTCAA 1200
Qy 1231 GTACTGATTTGCAAGTATGATGACCGAATGTTCCAGCTGAGAGACTATATCACTCT 1290
Db 1201 GTACTGATTTGCAAGTATGATGACCGAATGTTCCAGCTGAGAGACTATATCACTCT 1260
Qy 1291 GGAGGGAGGCGCTCTTATTTAGAGGCTCTACAGAAAACTTTGAGAAAGATAGCAAACT 1350
Db 1261 GGAGGGAGGCGCTCTTATTTAGAGGCTCTACAGAAAACTTTGAGAAAGATAGCAAACT 1320
Qy 1351 GAAGAGAGTCTGGATGAATTTAGAAAGAGACAGTCTAAAGAGAAATAGGAATCTCAC 1410
Db 1321 GAAGAGAGTCTGGATGAATTTAGAAAGAGACAGTCTAAAGAGAAATAGGAATCTCAC 1380
Qy 1411 AAAAGACCTAAAGACCCCAAAACAAAGAACGGGCTTTACCAAGAAAGAAACCTCAAA 1470
Db 1381 AAAAGACCTAAAGACCCCAAAACAAAGAACGGGCTTTACCAAGAAAGAAACCTCAAA 1440
Qy 1471 AAGAAAAAGGGAAGTATGACCTTAATCAATGATGAGAAAACTGAAAGAACTGGAAGAG 1530
Db 1441 AAGAAAAAGGGAAGTATGACCTTAATCAATGATGAGAAAACTGAAAGAACTGGAAGAG 1500
Qy 1531 GAAGGAGATGTCGAGGAAGATATGTCGAGAAATAGCAGTAGCCCGAAGAGTGCCTCG 1590
Db 1501 GAAGGAGATGTCGAGGAAGATATGTCGAGAAATAGCAGTAGCCCGAAGAGTGCCTCG 1560
Qy 1591 GAAGAAATTAAGCATGAAGATTTGATGTAATTTGTCATCGATGCTGCTTTGCGAATC 1650
Db 1561 GAAGAAATTAAGCATGAAGATTTGATGTAATTTGTCATCGATGCTGCTTTGCGAATC 1620
Qy 1651 CTGAAGAAAGCCCTCAGTATATCATCCGCTTCTGAGGCTGAGAGCCCTATATCTCTG 1710
Db 1621 CTGAAGAAAGCCCTCAGTATATCATCCGCTTCTGAGGCTGAGAGCCCTATATCTCTG 1680
Qy 1711 ACAAGGATCTACATGAAGTGAAGCAAGATGCTGTTCTTAAGAGCAGAGTAAAC 1770
Db 1681 ACAAGGATCTACATGAAGTGAAGCAAGATGCTGTTCTTAAGAGCAGAGTAAAC 1740
Qy 1771 TTTGTTGCGAGCTTGAATTTACAGGCGAGTAGGCTTGGAACCTCTGAGGGTTTAC 1830

Db 1741 TTTGTTGCGAGCTTGAATTTACAGGCGAGTAGGCGCTGGAAACCTGAGGGTTTAC 1800
Qy 1831 TTTCTTATATAGGAGGTTCACTGAGAGACACGCTATCTCACTGCTTTGCGGAAAGAA 1890
Db 1801 TTTCTTATATAGGAGGTTCACTGAGAGACACGCTATCTCACTGCTTTGCGGAAAGAA 1860
Qy 1891 AAGGAGCTTTTGAAGAACTGATAGGAAAAAGCAAGCATGTTTGTCCCTGGAAGAAAGA 1950
Db 1861 AAGGAGCTTTTGAAGAACTGATAGGAAAAAGCAAGCATGTTTGTCCCTGGAAGAAAGA 1920
Qy 1951 GAAGGAGAGATGAAGCAAACTTAGACCTAGTAAGAGGCAACAGCATCTGAGATGTTTCC 2010
Db 1921 GAAGGAGAGATGAAGCAAACTTAGACCTAGTAAGAGGCAACAGCATCTGAGATGTTTCC 1980
Qy 2011 ACTGACACTCGGAAAGCCGGTGGCCAGAGACAGAAATGCTACACAGCAAGCATAGTTGTG 2070
Db 1981 ACTGACACTCGGAAAGCCGGTGGCCAGAGACAGAAATGCTACACAGCAAGCATAGTTGTG 2040
Qy 2071 GATATGCGTGAATTTGGAAGTGAAGTCCATCTCTGATCCATGCTGCGAGCATGACATT 2130
Db 2041 GATATGCGTGAATTTGGAAGTGAAGTCCATCTCTGATCCATGCTGCGAGCATGACATT 2100
Qy 2131 GAACCCGTGACTTTAGAGGTTGAGATTTACATCTCACTCCAGAAATGCTGCTGAGAGCCG 2190
Db 2101 GAACCCGTGACTTTAGAGGTTGAGATTTACATCTCACTCCAGAAATGCTGCTGAGAGCCG 2160
Qy 2191 AAGATATCATGATTTAATGGGCTTTAATTAAGGCGCGCTCTACAGCAGATGATC 2250
Db 2161 AAGATATCATGATTTAATGGGCTTTAATTAAGGCGCGCTCTACAGCAGATGATC 2220
Qy 2251 TCCATGTCGCCGCTACTACAGCGTCCGCTGCTTCTGATTTGATTTGAACCTAGCAAGCT 2310
Db 2221 TCCATGTCGCCGCTACTACAGCGTCCGCTGCTTCTGATTTGATTTGAACCTAGCAAGCT 2280
Qy 2311 TTTCTCTCACTTCCCGAGGTCCTGTTTTCAGAGATCTCAGAGATGACATTAATGTTCC 2370
Db 2281 TTTCTCTCACTTCCCGAGGTCCTGTTTTCAGAGATCTCAGAGATGACATTAATGTTCC 2340
Qy 2371 AAATCTCACTTTCTACACTTCACTTCCCGAGATGAGATTTCTGTTGCGCCCTCTCT 2430
Db 2341 AAATCTCACTTTCTACACTTCACTTCCCGAGATGAGATTTCTGTTGCGCCCTCTCT 2400
Qy 2431 CATGCAAGGCGGAGTCTTTGAGAGGCTGAACCAAGCAAGCCACGCTGATGCGGCG 2490
Db 2401 CATGCAAGGCGGAGTCTTTGAGAGGCTGAACCAAGCAAGCCACGCTGATGCGGCG 2460
Qy 2491 ACAGCACTGCGCATTTACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCT 2550
Db 2461 ACAGCACTGCGCATTTACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCT 2520
Qy 2551 GGTCCCAAGACTTCTGTTAAATGCAAGGAGGATGATGCCAAAGTGGCGGCTCTTG 2610
Db 2521 GGTCCCAAGACTTCTGTTAAATGCAAGGAGGATGATGCCAAAGTGGCGGCTCTTG 2580
Qy 2611 ATGCACACGTTTAAAGACATGACAGATTTAGCAGGCTGTCAAGAGAGAGCTCAAGAT 2670
Db 2581 ATGCACACGTTTAAAGACATGACAGATTTAGCAGGCTGTCAAGAGAGAGCTCAAGAT 2640
Qy 2671 ATTCTGGGGAATGTCGAAATGCCAAGACGTTTATGATTTCACTCAACCTTTTGCA 2730
Db 2641 ATTCTGGGGAATGTCGAAATGCCAAGACGTTTATGATTTCACTCAACCTTTTGCA 2700
Qy 2731 GAAGTGTATCAAAAGGAAAGGAAAAAGTGAAGAGAGGCTGTTTCTTATCCCAT 2790
Db 2701 GAAGTGTATCAAAAGGAAAGGAAAAAGTGAAGAGAGGCTGTTTCTTATCCCAT 2760
Qy 2791 GCTTACTTTTTCAGGCGCTCTTCCAGACATCAATAGTCAATTTATTAATTTGTTTG 2850
Db 2761 GCTTACTTTTTCAGGCGCTCTTCCAGACATCAATAGTCAATTTATTAATTTGTTTG 2820
Qy 2851 CTATTTCATCTTTTTCATATGCTTTAATGATTTGACGTGAGCAGAGCCAG 2904
Db 2821 CTATTTCATCTTTTTCATATGCTTTAATGATTTGACGTGAGCAGAGCTTAC 2874

RESULT 3
ARI83109
LOCUS ARI83109 2718 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 14 from patent US 6340566.
ACCESSION ARI83109
VERSION ARI83109.1 GI:20226702
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2718)
AUTHORS McCutchen-Maloney,S.L.
TITLE Detection and quantitation of single nucleotide polymorphisms, DNA
sequence variations, DNA mutations, DNA damage and DNA mismatches
JOURNAL Patent: US 6340566-A 14 22-JAN-2002;
FEATURES
source 1..2718
Location/Qualifiers
BASE COUNT 827 a 588 c 634 g 669 t
ORIGIN
Query Match 70.5%; Score 2718; DB 6; Length 2718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 ATGGCCCCGCTGCTGAGTACGAGCAGACGCTGCTGAACTGCTCGACACTGACGG 105
DB 1 ATGGCCCCGCTGCTGAGTACGAGCAGACGCTGCTGAACTGCTCGACACTGACGG 60
QY 106 CTAGTAGTGTGCGCCCGCGGGCTCGCGCGGACCGGCTCCTACCACTTCTCCAGCT 165
DB 61 CTAGTAGTGTGCGCCCGCGGGCTCGCGCGGACCGGCTCCTACCACTTCTCCAGCT 120
QY 166 CACTGCCACCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
DB 121 CACTGCCACCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 226 TTATCAATCAGCTGAGAGATAGAAAGATTGAACACCTCCGCGCGGTGAACAAATGA 285
DB 181 TTATCAATCAGCTGAGAGATAGAAAGATTGAACACCTCCGCGCGGTGAACAAATGA 240
QY 286 ATCACAAGCAACAGTGCCTATGAAGTTTACACACAAGGTGTGTTATTTGGCACAAGT 345
DB 241 ATCACAAGCAACAGTGCCTATGAAGTTTACACACAAGGTGTGTTATTTGGCACAAGT 300
QY 346 AGGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 405
DB 301 AGGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 406 TTGGTGTATAGAGCCCAAGAAATATCGAGTCTTGTCAAGAACATTCATCTTGGCCCTC 465
DB 361 TTGGTGTATAGAGCCCAAGAAATATCGAGTCTTGTCAAGAACATTCATCTTGGCCCTC 420
QY 466 TTTCCGCCAGAAAACAACAGCTGTTTATTAAGCTTTCACAGACAACTGCTGTGCTT 525
DB 421 TTTCCGCCAGAAAACAACAGCTGTTTATTAAGCTTTCACAGACAACTGCTGTGCTT 480
QY 526 GATACGTGTTTGTGATGTGGAAGAGATGATGAGAAATCTTTTGTGAGGAACCTGAT 585
DB 481 GATACGTGTTTGTGATGTGGAAGAGATGAGAAATCTTTTGTGAGGAACCTGAT 540
QY 586 CTGTGGCCAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
DB 541 CTGTGGCCAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 646 GTAGAAATCCATGTTTCTATGACACCTACCATGCTTGTGATACAGACTCTATCTAGAC 705
DB 601 GTAGAAATCCATGTTTCTATGACACCTACCATGCTTGTGATACAGACTCTATCTAGAC 660
QY 706 ATTTTAATGCAATGCTTAAGAACTAAATGCCATACCCATCGTTGAAGTGAAGAT 765
DB 1 ATTTTAATGCAATGCTTAAGAACTAAATGCCATACCCATCGTTGAAGTGAAGAT 1

DB 661 ATTTTAATGCAATGCTTAAGAACTAAATGCCATACCCATCGTTGAAGTGAAGAT 720
QY 766 TTATCTTTAGAAAATGCTATTTGGAACCTTTTGACAGCAATCCGCCATTATCTGGAT 825
DB 721 TTATCTTTAGAAAATGCTATTTGGAACCTTTTGACAGCAATCCGCCATTATCTGGAT 780
QY 826 CTTTGTGGCAGCAGCTTGGAGCCAAAGACTAAATCCCTTGTGAGCAATTTGAAGATTA 885
DB 781 CTTTGTGGCAGCAGCTTGGAGCCAAAGACTAAATCCCTTGTGAGCAATTTGAAGATTA 840
QY 886 CGAATCTTGTGAGCAGTATCTCTCAGTATGATGATGATGATGATGATGATGATGATGAT 945
DB 841 CGAATCTTGTGAGCAGTATCTCTCAGTATGATGATGATGATGATGATGATGATGATGAT 900
QY 946 TCTCTGAGCAACGCAAAAGCTTTTGTGAGCAATTTGAGCTGCTGCTTCTGACTCC 1005
DB 901 TCTCTGAGCAACGCAAAAGCTTTTGTGAGCAATTTGAGCTGCTGCTTCTGACTCC 960
QY 1006 AGCAGCTCGATGTTTAAATGCTGAGCAAGGCTTATCATCTTCACATGCCAAATG 1065
DB 961 AGCAGCTCGATGTTTAAATGCTGAGCAAGGCTTATCATCTTCAGATGCCAAATG 1020
QY 1066 AGTAAAAAGAAAAATATCTGAAAAATGAAATTAAGAAAGGAGAGAAACAAAAAG 1125
DB 1021 AGTAAAAAGAAAAATATCTGAAAAATGAAATTAAGAAAGGAGAGAAACAAAAAG 1080
QY 1126 GAACTGCTCTAGAAACCAACCAAGTGGAGGCACTGACTGATATTAAGAATAAT 1185
DB 1081 GAACTGCTCTAGAAACCAACCAAGTGGAGGCACTGACTGATATTAAGAATAAT 1140
QY 1186 GAGCAGAAAATATGAGAGATGAGAGCTTGTGGGTCAGAGTCACTACTGATTTTGA 1245
DB 1141 GAGCAGAAAATATGAGAGATGAGAGCTTGTGGGTCAGAGTCACTACTGATTTTGA 1200
QY 1246 AGTATGACGGAACATGTTCCAGCTGAGAGACTATATCATCTTGTGAGCGAGGCTTC 1305
DB 1201 AGTATGACGGAACATGTTCCAGCTGAGAGACTATATCATCTTGTGAGCGAGGCTTC 1260
QY 1306 TTATGAGGCTCTACAGAAAACCTTTGAGAAAGATAGCAAAAGCTGATGATGATG 1365
DB 1261 TTATGAGGCTCTACAGAAAACCTTTGAGAAAGATAGCAAAAGCTGATGATGATG 1320
QY 1366 AAATTTAGAGAGAGAGACAGTTCACAAAGAAATAGAGAAATCTCACAAAGCTTAAGAC 1425
DB 1321 AAATTTAGAGAGAGAGACAGTTCACAAAGAAATAGAGAAATCTCACAAAGCTTAAGAC 1380
QY 1426 CCCCAGAAACAAAGAGAGGCTTCTACCAAGAAAGAACCTCTCAAAAGAAAAAGGAG 1485
DB 1381 CCCCAGAAACAAAGAGAGGCTTCTACCAAGAAAGAACCTCTCAAAAGAAAAAGGAG 1440
QY 1486 TTGACCTTAACCTCAATGTGAGAGAAACCTGAAAGATGAGAGAGAGATGTCAG 1545
DB 1441 TTGACCTTAACCTCAATGTGAGAGAAACCTGAAAGATGAGAGAGAGATGTCAG 1500
QY 1546 GAAGATATCTGAGAGAAATAGCAGTAGCCAGAAAGCTGCCGAGAGAAATTAAGAT 1605
DB 1501 GAAGATATCTGAGAGAAATAGCAGTAGCCAGAAAGCTGCCGAGAGAAATTAAGAT 1560
QY 1606 GAAAGATTTGATGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGAT 1665
DB 1561 GAAAGATTTGATGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1666 ACTATCATCATCCGCTTGTGAGGTCAGAGACCCCTATGCTGTCAGCAAGGATACAT 1725
DB 1621 ACTATCATCATCCGCTTGTGAGGTCAGAGACCCCTATGCTGTCAGCAAGGATACAT 1680
QY 1726 GAAAGTGAACCAAGATACCTGCTTTATGACGAGAGACTAACCCTTTGCTGGAGCTT 1785
DB 1681 GAAAGTGAACCAAGATACCTGCTTTATGACGAGAGACTAACCCTTTGCTGGAGCTT 1740
QY 1786 GAAATTTACAGGCGAGTATGAGGCTGGGAAACCTCTGAGGCTTACCTTTCTATACGGA 1845
DB 1741 GAAATTTACAGGCGAGTATGAGGCTGGGAAACCTCTGAGGCTTACCTTTCTATACGGA 1800

QY 1846 GGTCAACTGAGAACAAAGCTATCTCACTGCTTGGGAAAGAAAGAAAGCTTTTGA 1905
|||||
DB 1801 GGTCAACGAGAGAACAAAGCTATCTCACTGCTTGGGAAAGAAAGAAAGCTTTTGA 1860
QY 1906 AACTCATAGAGGAAAGAAAGCAAGATGTTGCTCCCTGAAGAAAGAAAGCAAGATGAA 1965
|||||
DB 1861 AACTCATAGAGGAAAGAAAGCAAGATGTTGCTCCCTGAAGAAAGAAAGCAAGATGAA 1920
QY 1966 ACAACTTAGACCTAGTAGAGGACAGATCTGAGATGTTTCCACTGACACTGGGAA 2025
|||||
DB 1921 ACAAACTTAGACCTAGTAGAGGACAGATCTGAGATGTTTCCACTGACACTGGGAA 1980
QY 2026 GCCGGTGGCCAGAGAACAGATGTTACACAGCAAGCAATAGTGGATATGCTGAATTT 2085
|||||
DB 1981 GCCGGTGGCCAGAGAACAGATGTTACACAGCAAGCAATAGTGGATATGCTGAATTT 2040
QY 2086 CGAAGTGAAGCTTCATCTCTGATCCATGCTGGGGCATTTGACATTTGAACCCGCTGATTTA 2145
|||||
DB 2041 CGAAGTGAAGCTTCATCTCTGATCCATGCTGGGGCATTTGACATTTGAACCCGCTGATTTA 2100
QY 2146 GAGGTGGGATTTACATCTCTCACTCCAGAAATGTGGGTGGAGCGGAGATATCAAGTGTAT 2205
|||||
DB 2101 GAGGTGGGATTTACATCTCTCACTCCAGAAATGTGGGTGGAGCGGAGATATCAAGTGTAT 2160
QY 2206 TTAATCGGCTCTTTAAATTAACGGCGCTCTACAGCAAGTGCATCTCATATGCCGCTAC 2265
|||||
DB 2161 TTAATCGGCTCTTTAAATTAACGGCGCTCTACAGCAAGTGCATCTCATATGCCGCTAC 2220
QY 2266 TACAAGCGTCCGCTGCTTGTATTTGAATTTGACCTAGCAAGCCTTTCTCTCACTTCC 2325
|||||
DB 2221 TACAAGCGTCCGCTGCTTGTATTTGATTTGACCTAGCAAGCCTTTCTCTCACTTCC 2280
QY 2326 CGAGGTGCTTGTTCAGAGATCTCCAGCATGATAGTATCCAACTCCTCTCTT 2385
|||||
DB 2281 CGAGGTGCTTGTTCAGAGATCTCCAGCATGATAGTATCCAACTCCTCTCTT 2340
QY 2386 ACACCTTCACTTCCAGAGATGATCTCTGTCGCCCTCTCATATGCAAGCGGAG 2445
|||||
DB 2341 ACACCTTCACTTCCAGAGATGATCTCTGTCGCCCTCTCATATGCAAGCGGAG 2400
QY 2446 TTGTTTGAAGAGCTGAAACCAAGCAAGCAAGCCTGATGCGGCGAGCACTGSCATT 2505
|||||
DB 2401 TTGTTTGAAGAGCTGAAACCAAGCAAGCAAGCCTGATGCGGCGAGCACTGSCATT 2460
QY 2506 AAGCAGATTTTGAAACCTTCCCGAGTACAGAAATATATCTCTGTCGCCCAAGCTTC 2565
|||||
DB 2461 AAGCAGATTTTGAAACCTTCCCGAGTACAGAAATATATCTCTGTCGCCCAAGCTTC 2520
QY 2566 TTGTTTAAATGCGAGGGGTGAATGCAAAACTGCCGCTCTGATGACCAACGTTAAG 2625
|||||
DB 2521 TTGTTTAAATGCGAGGGGTGAATGCAAAACTGCCGCTCTGATGACCAACGTTAAG 2580
QY 2626 AACATCGCAGAAATTTAGCAGCCTGTCTACAAGAGAGAGCTCAGAGATTTCTGGGGAATGCT 2685
|||||
DB 2581 AACATCGCAGAAATTTAGCAGCCTGTCTACAAGAGAGAGCTCAGAGATTTCTGGGGAATGCT 2640
QY 2686 GGAATATGCAAAAGCTTTATGATTTTCACTCACTCTTTTGCAGAAAGTCTGATCAAAA 2745
|||||
DB 2641 GGAATATGCAAAAGCTTTATGATTTTCACTCACTCTTTTGCAGAAAGTCTGATCAAAA 2700
QY 2746 GGAAGAGGAGAAAGTGA 2763
|||||
DB 2701 GGAAGAGGAGAAAGTGA 2718

RESULT 4
AR203246 2718 bp DNA linear PAT 20-JUN-2002
LOCUS AR203246
DEFINITION Sequence 14 from patent US 6365355.
ACCESSION AR203246
VERSION AR203246.1 GI:21499585
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2718)
AUTHORS McCutchen-Maloney,S.L.
TITLE Chimeric proteins for detection and quantitation of DNA mutations,
JOURNAL DNA sequence variations, DNA damage and DNA mismatches
FEATURES
source location/Qualifiers
BASE COUNT 827 a 588 c 634 g 669 t
ORIGIN
Query Match 70.5%; Score 2718; DB 6; Length 2718;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 ATGGCGCCGCTGCTGAGTAGAGCAGCAGCTGCTGGAACCTCTCGACACTGACGG 105
|||||
DB 1 ATGGCGCCGCTGCTGAGTAGAGCAGCAGCTGCTGGAACCTCTCGACACTGACGG 60
QY 106 CTAGTAGTGTGGCGCGCGGCGCTGGCGGACCGGCTCTTACACTTCTCCAGCTG 165
|||||
DB 61 CTAGTAGTGTGGCGCGCGGCGCTGGCGGACCGGCTCTTACACTTCTCCAGCTG 120
QY 166 CACTGCCACCCAGCCCTGCTGCTGCTGCTCAACACGACGCGCCGAGAGAGATAT 225
|||||
DB 121 CACTGCCACCCAGCCCTGCTGCTGCTGCTCAACACGACGCGCCGAGAGAGATAT 180
QY 226 TTTATCAATCACTGAAATAGAGAGATTTGAACACCTCCCTGCGCTTAACAAATGAA 285
|||||
DB 181 TTTATCAATCACTGAAATAGAGAGATTTGAACACCTCCCTGCGCTTAACAAATGAA 240
QY 286 ATCAACAGCAACAGTGCCTATGAATTTACACAAAGTGTGTATATTTGGCAAAAT 345
|||||
DB 241 ATCAACAGCAACAGTGCCTATGAATTTACACAAAGTGTGTATATTTGGCAAAAT 300
QY 346 AGGATACCTGTGGTGTACTCTTGAATGATGATGATGATGATGATGATGATGATGAT 405
|||||
DB 301 AGGATACCTGTGGTGTACTCTTGAATGATGATGATGATGATGATGATGATGATGAT 360
QY 406 TTGTTTATAGAGCCCAAGAAATATGAGTCTTGTCAAGAAAGCAATCTTTCGCCCTC 465
|||||
DB 361 TTGTTTATAGAGCCCAAGAAATATGAGTCTTGTCAAGAAAGCAATCTTTCGCCCTC 420
QY 466 TTTCCGCAAGAAACAAAGCTGTTTATTAAGCTTTTACAGACATGCTGTGCTTT 525
|||||
DB 421 TTTCCGCAAGAAACAAAGCTGTTTATTAAGCTTTTACAGACATGCTGTGCTTT 480
QY 526 GATACGTGTTTGTGATGGAAGAGTATGAGAAATCTTTTGTGAGGAATGATAT 585
|||||
DB 481 GATACGTGTTTGTGATGGAAGAGTATGAGAAATCTTTTGTGAGGAATGATAT 540
QY 586 CTGTGGCAAGGTTCCATGTAGCAGTAACCTATTTTGAACACCAACCTGAAGTT 645
|||||
DB 541 CTGTGGCAAGGTTCCATGTAGCAGTAACCTATTTTGAACACCAACCTGAAGTT 600
QY 646 GTAGAAATCCATGTTTCTATGACACCTACCATGCTGTATACAGCTCTATACGAG 705
|||||
DB 601 GTAGAAATCCATGTTTCTATGACACCTACCATGCTGTATACAGCTCTATACGAG 660
QY 706 ATTTTAAATGATGCTAAGAAAGCACTAAATGCAATCCATCCCTGGAAGTGAAGAT 765
|||||
DB 661 ATTTTAAATGATGCTAAGAAAGCACTAAATGCAATCCATCCCTGGAAGTGAAGAT 720
QY 766 TTAATCTTGAAGAAATGCTATTTGAAAGCTTTTGAACAGCAATCCCATTAATCGAT 825
|||||
DB 721 TTAATCTTGAAGAAATGCTATTTGAAAGCTTTTGAACAGCAATCCCATTAATCGAT 780
QY 826 CCTTTGTGACACAGCTTGAGGCAAGCACTAAATCTTATGTTAGATTTGAAGATATTA 885
|||||
DB 781 CCTTTGTGACACAGCTTGAGGCAAGCACTAAATCTTATGTTAGATTTGAAGATATTA 840

QY 886 CGAATCTTGCAGTAATCTCTCAGATATGATGTGTGACATTTCTTAATCTTCTGAA 945
 Db 841 CGAACTTTGGCTGAGATATCTCTCAGATATGATGTGTGACATTTCTTAATCTTCTGAA 900
 QY 946 TCTCTGAGAGCAAGCAAAAGCTTTTGTGAGATTCAGGTGGCTTTCTTGACACC 1005
 Db 901 TCTCTGAGAGCAAGCAAAAGCTTTTGTGAGATTCAGGTGGCTTTCTTGACACC 960
 QY 1006 AGCACTTCGATGTTTATTAATGCTCGAGCAAGGCTTATCATCTTCCAGATGCCAAATG 1065
 Db 961 AGCACTTCGATGTTTATTAATGCTCGAGCAAGGCTTATCATCTTCCAGATGCCAAATG 1020
 QY 1066 AGTAAAAAGAAAAATATCTGAAAAATGAAATTAAGAGAGGGAAGAAACAAAAAG 1125
 Db 1021 AGTAAAAAGAAAAATATCTGAAAAATGAAATTAAGAGAGGGAAGAAACAAAAAG 1080
 QY 1126 GAATGCTCTAGAAAGCAACCCAAAGTGGAGGCACTGACTGATGAATTTAAAGAAAT 1185
 Db 1081 GAATGCTCTAGAAAGCAACCCAAAGTGGAGGCACTGACTGATGAATTTAAAGAAAT 1140
 QY 1186 GAGCAGAAAAATAGAGAGATGAAAGCTTTGGTCCAGTCAAGTACGATTTGTGCA 1245
 Db 1141 GAGCAGAAAAATAGAGAGATGAAAGCTTTGGTCCAGTCAAGTACGATTTGTGCA 1200
 QY 1246 AGTATGACGCAACATGTTCCAGCTGAGAGACTATATCACTCTTGGAGGCGAGCCCTTC 1305
 Db 1201 AGTATGACGCAACATGTTCCAGCTGAGAGACTATATCACTCTTGGAGGCGAGCCCTTC 1260
 QY 1306 TTATTTAGGCTCTACAGAAAAACCTTTGAGAGAGATAGCAAAAGCTGAAGATGTGATG 1365
 Db 1361 TTATTTAGGCTCTACAGAAAAACCTTTGAGAGAGATAGCAAAAGCTGAAGATGTGATG 1320
 QY 1366 AAATTTAGGAAGAAAGCAGTTCAAGAGAAATAGAAATCTCAAAAGACCTAAAGAC 1425
 Db 1321 AAATTTAGGAAGAAAGCAGTTCAAGAGAAATAGAAATCTCAAAAGACCTAAAGAC 1380
 QY 1426 CCCCCAAAACAAGAGGGGCTTTACCAAGAAAGAAACCCCTCAAAAAGAAAAAGCGAG 1485
 Db 1381 CCCCCAAAACAAGAGGGGCTTTACCAAGAAAGAAACCCCTCAAAAAGAAAAAGCGAG 1440
 QY 1486 TTGACCTTAATCTCAATGTGTAGAAAGAACTGGAAGAGGAAAGAGATGTGAG 1545
 Db 1441 TTGACCTTAATCTCAATGTGTAGAAAGAACTGGAAGAGGAAAGAGATGTGAG 1500
 QY 1546 GAAGGATATGCTGAGAAATAGCAGTACCCAGAAAGCTGCCGGAAGAAATTAACAT 1605
 Db 1501 GAAGGATATGCTGAGAAATAGCAGTACCCAGAAAGCTGCCGGAAGAAATTAACAT 1560
 QY 1606 GAAAGATTTGATGTAATTTGTCTATCGGATGTGCTTTGGAAATCTGAAAGAACCCCTC 1665
 Db 1561 GAAAGATTTGATGTAATTTGTCTATCGGATGTGCTTTGGAAATCTGAAAGAACCCCTC 1620
 QY 1666 ACTATCATCATCCGCTTTCGGGTGTGAGAGAGACCCCTATGCTGTGAAAGAGGTACTCAT 1725
 Db 1621 ACTATCATCATCCGCTTTCGGGTGTGAGAGAGACCCCTATGCTGTGAAAGAGGTACTCAT 1680
 QY 1726 GAAGTGGAGCAAGATACGTGTCTTTATGACAGAGCTTAACCTTTGTTGGCAGCTT 1785
 Db 1681 GAAGTGGAGCAAGATACGTGTCTTTATGACAGAGCTTAACCTTTGTTGGCAGCTT 1740
 QY 1786 GAAATTTACAGAGGCGATAGGCTGGGAAACCTCTGAGGGTTTACTTTCTATATACGA 1845
 Db 1741 GAAATTTACAGAGGCGATAGGCTGGGAAACCTCTGAGGGTTTACTTTCTATATACGA 1800
 QY 1846 GCTTCAACTGAGCAACACGCTATCTACGCTTTGGGGAAGAAAGAAAGCTTTTGA 1905
 Db 1801 GCTTCAACTGAGCAACACGCTATCTACGCTTTGGGGAAGAAAGAAAGCTTTTGA 1860
 QY 1906 AAATCTAATAGGGAAGAAAGCAAGCATGTTGCCCTGAAGAAAGAGAGGAGAGATGAA 1965
 Db 1861 AAATCTAATAGGGAAGAAAGCAAGCATGTTGCCCTGAAGAAAGAGAGGAGAGATGAA 1920

QY 1966 ACAACTTATGACCTATGTAAGAGGACAGCATCTGCAATGTTTCCACTTGACACTCGGAA 2025
 Db 1921 ACAACTTATGACCTATGTAAGAGGACAGCATCTGCAATGTTTCCACTTGACACTCGGAA 1980
 QY 2026 GCCGCTGGCCAGAAACAGAAATGATACAGCAAGACATAGTTTGGCAATATCGGAAATTT 2085
 Db 1981 GCCGCTGGCCAGAAACAGAAATGATACAGCAAGACATAGTTTGGCAATATCGGAAATTT 2040
 QY 2086 CGAATGAGCTTCCATCTCTGATATCAGTGTGCGGCAATGACATTTCAACCCGTACTTTA 2145
 Db 2041 CGAATGAGCTTCCATCTCTGATATCAGTGTGCGGCAATGACATTTGAACCCGTACTTTA 2100
 QY 2146 GAGGTGAGATTTACATCCCTACCTCCAGAAATGCGCGGAGCGCCAGAGATAGAT 2205
 Db 2101 GAGGTGAGATTTACATCCCTACCTCCAGAAATGCGCGGAGCGCCAGAGATAGAT 2160
 QY 2206 TTATTCGCTCTTTTAAATTAACGGCCGCTCTACAGCCAGTGCATCTCCATGCTCCGCTAC 2265
 Db 2161 TTATTCGCTCTTTTAAATTAACGGCCGCTCTACAGCCAGTGCATCTCCATGCTCCGCTAC 2220
 QY 2266 TACAAGGCTCCGCTGCTTCTGATTTGAGTTTGAACCTTACCAAGCCTTTCTCTCACTTCC 2325
 Db 2221 TACAAGGCTCCGCTGCTTCTGATTTGAGTTTGAACCTTACCAAGCCTTTCTCTCACTTCC 2280
 QY 2326 CGAGTGCCTTGTTCAGAGATCTCCAGCAATGACATTTAGTTTCAAACTCACTCTT 2385
 Db 2281 CGAGTGCCTTGTTCAGAGATCTCCAGCAATGACATTTAGTTTCAAACTCACTCTT 2340
 QY 2386 ACACCTTCACTTCCCGACACTACGAGATTTCTGTGTCCTCTCTCTCAATGCAAGCGCGAG 2445
 Db 2341 ACACCTTCACTTCCCGACACTACGAGATTTCTGTGTCCTCTCTCTCAATGCAAGCGCGAG 2400
 QY 2446 TTGTTTGAAGAGCTGAAGAAAGCAAGCAAGCAAGCAAGCTATGGGAGACAGACTGGCAT 2505
 Db 2401 TTGTTTGAAGAGCTGAAGAAAGCAAGCAAGCAAGCTATGGGAGACAGACTGGCAT 2460
 QY 2506 ACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATATCTCTCTCCCAAGACTTC 2565
 Db 2461 ACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATATCTCTCTCCCAAGACTTC 2520
 QY 2566 TTGTTTAAATATGCCAGGGGTGAATGCCAAAAAATGCGGCTCTTGAATGACACAGCTTAA 2625
 Db 2521 TTGTTTAAATATGCCAGGGGTGAATGCCAAAAAATGCGGCTCTTGAATGACACAGCTTAA 2580
 QY 2626 AACATCGCAGATTTAGAGCCCTGTCTCAAGAGAGCTCACAGATTTTCGGGAAATGCT 2685
 Db 2581 AACATCGCAGATTTAGAGCCCTGTCTCAAGAGAGCTCACAGATTTTCGGGAAATGCT 2640
 QY 2686 GCAAAATGCCAAACAGCTTTATGATTTTACACCTCTTTTGCAGAAAGTCTGATCAAAA 2745
 Db 2641 GCAAAATGCCAAACAGCTTTATGATTTTACACCTCTTTTGCAGAAAGTCTGATCAAAA 2700
 QY 2746 GGAAGAGGGAAGAAAGTGA 2763
 Db 2701 GGAAGAGGGAAGAAAGTGA 2718

RESULT 5
 AX277006
 LOCUS AX277006 2718 bp DNA linear PAT 29-OCT-2001
 DEFINITION Sequence 14 from Patent WO0173079.
 ACCESSION AX277006
 VERSION AX277006.1 GI:16548673
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 McCutchen-Maloney, S.L.
 TITLE Chimeric proteins for detection and quantitation of dna mutations,
 dna sequence variations, dna damage and dna mismatches
 JOURNAL Patent: WO 0173079-A 14 04-OCT-2001;

The Regents of The University of California (US)

Location/Qualifiers

1..2718

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 827 a 588 c 634 g 669 t

ORIGIN

Query Match 70.5%; Score 2718; DB 6; Length 2718;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

46 ATGGCGCCGCTCTGGAGTACGACGACACTGCTGGAAGTCTGCACACTGACGG 105
1 ATGGCGCCGCTCTGGAGTACGACGACACTGCTGGAAGTCTGCACACTGACGG 60
106 CTAGTAGTGTGGCGCGCGGCTCGCGCGGACCGGCTCCTTACACTTTTCCAGCTG 165
61 CTAGTAGTGTGGCGCGCGGCTCGCGCGGACCGGCTCCTTACACTTTTCCAGCTG 120
166 CACTGCACCCAGCCCTGCTGCTGCTGCTCAACAGCAGCGCGGCGAGGAGGAT 225
121 CACTGCACCCAGCCCTGCTGCTGCTGCTCAACAGCAGCGCGGCGAGGAGGAT 180
226 TTTATCAATCAGCTGAAGTAGAAGAGTGAACACCTCCCTCGCGCTGAACAAATGA 285
181 TTTATCAATCAGCTGAAGTAGAAGAGTGAACACCTCCCTCGCGCTGAACAAATGA 240
286 ATCACAGCAGCAGTCTGATGAAGTTTACACAAAGTGTGTATATTTGCGACAAGT 345
241 ATCACAGCAGCAGTCTGATGAAGTTTACACAAAGTGTGTATATTTGCGACAAGT 300
-346 AGATACCTGTGGTGAATCTTGAATACCTTGAATACCTTGAATACCTTGAATAC 405
301 AGATACCTGTGGTGAATCTTGAATACCTTGAATACCTTGAATACCTTGAATAC 360
406 TTGCTGTATAGAGCCACAGAAATGAGTCTTGTCAAGAGCAATCATCTTGCCTC 465
361 TTGCTGTATAGAGCCACAGAAATGAGTCTTGTCAAGAGCAATCATCTTGCCTC 420
466 TTTCCGCGAGAAAACAAAGCTGTTTATTTAAAGCTTTACAGACATCTGTTGCTTT 525
421 TTTCCGCGAGAAAACAAAGCTGTTTATTTAAAGCTTTACAGACATCTGTTGCTTT 480
526 GATACGCTTTTGTGATGTGGAAGAGTGAAGAAATCTTTGTGAGGAACCTGAT 585
481 GATACGCTTTTGTGATGTGGAAGAGTGAAGAAATCTTTGTGAGGAACCTGAT 540
586 CTGTGGCCAGAGTTCATGATAGCAGTAACTATTTTGAACAGCAGCAAACTGAAGT 645
541 CTGTGGCCAGAGTTCATGATAGCAGTAACTATTTTGAACAGCAGCAAACTGAAGT 600
646 GTAGAAATTCATGTTTATGACACTACCATGCTTGTATACAGCTCTTACTGAC 705
601 GTAGAAATTCATGTTTATGACACTACCATGCTTGTATACAGCTCTTACTGAC 660
706 ATTTAAATGCATGCTTAAGAGACTAAATATCCATACCCATCGCTTAAAGTGAAGAT 765
661 ATTTAAATGCATGCTTAAGAGACTAAATATCCATACCCATCGCTTAAAGTGAAGAT 720
766 TTTATCTTGAAGAAATGCTATTGGAAGAACTTTTGAACAGCAATCGCCATTAATCGAT 825
721 TTTATCTTGAAGAAATGCTATTGGAAGAACTTTTGAACAGCAATCGCCATTAATCGAT 780
826 CTTTGTGGCAGCAGCTTGGAGCCAAAGACTAATCTTATGTTCAAGATTTGAAGATTA 885
781 CTTTGTGGCAGCAGCTTGGAGCCAAAGACTAATCTTATGTTCAAGATTTGAAGATTA 840
886 CGAATTTGCTGAGATATCTCTCAGATGATGATGTCACATTTCTTATATCTTGGAA 945
841 CGAATTTGCTGAGATATCTCTCAGATGATGATGTCACATTTCTTATATCTTGGAA 900
946 TCTGTGAGAGCAAGCAAAAGCTTTTGTGATGAATTCAGTTGGCTGTCTTGTGACTCC 1005

|||||
Db 901 TCTCTGAGAGCAACGAAAAAGCTTTTGGTACAAATTCAGGTGGCTGTTTGTGACTCC 960
1006 AGCACTTCATGTTTAAATGCTGCAGCAAGGCTTATCATCTTCCAGATGCCAAAATG 1065
961 AGCACTTCATGTTTAAATGCTGCAGCAAGGCTTATCATCTTCCAGATGCCAAAATG 1020
1066 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAGGGGAGCAACAAAAAG 1125
1021 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAGGGGAGCAACAAAAAG 1080
1126 GAACCTGCTCTAGAAAAGCAACCCAAAGTGGAGGCACTGATGATTTAAAGAAAT 1185
1081 GAACCTGCTCTAGAAAAGCAACCCAAAGTGGAGGCACTGATGATTTAAAGAAAT 1140
1186 GAGGCAAAATTAAGAGAGTGAAGTCTTGTGTGTCAGGTCAATTAATGATTTGCA 1245
1141 GAGGCAAAATTAAGAGAGTGAAGTCTTGTGTGTCAGGTCAATTAATGATTTGCA 1200
1246 AGTATGAGCGAATGTTCCAGCTGAGAGACTATATCATCTTGGAGGAGGCTTC 1305
1201 AGTATGAGCGAATGTTCCAGCTGAGAGACTATATCATCTTGGAGGAGGCTTC 1260
1306 TATTGAGGCTCTACAGAAAACCTTTGAGAAAGATAGCAAAAGCTGAGATG 1365
1261 TATTGAGGCTCTACAGAAAACCTTTGAGAAAGATAGCAAAAGCTGAGATG 1320
1366 AAATTTAGAGAGAGACAGTTCAAAAGATTTAGAAATCTCACAAAAGACTTAAGAC 1425
1321 AAATTTAGAGAGAGACAGTTCAAAAGATTTAGAAATCTCACAAAAGACTTAAGAC 1380
1426 CCCCAGAAAGAAAGGGCTTCTACAAAAGAAAGACCTCAAAAAGAAAACGGAAG 1485
1381 CCCCAGAAAGAAAGGGCTTCTACAAAAGAAAGACCTCAAAAAGAAAACGGAAG 1440
1486 TTGACCTTAATCTCAATGTTAGAGAAAACCTGAAAGACTGGAAGAGTGTGAG 1545
1441 TTGACCTTAATCTCAATGTTAGAGAAAACCTGAAAGACTGGAAGAGTGTGAG 1500
1546 GAAGATATGCTGAGAAATTAAGCACTAGCCAGAAAGCTGCCGCAAGAAATTAAGCAT 1605
1501 GAAGATATGCTGAGAAATTAAGCACTAGCCAGAAAGCTGCCGCAAGAAATTAAGCAT 1560
1606 GAAGATTTGATGTAATTTGATGATGAGTGGCTTGGAAATCTGAAAGAACCCCTC 1665
1561 GAAGATTTGATGTAATTTGATGATGAGTGGCTTGGAAATCTGAAAGAACCCCTC 1620
1666 ACTATCATCATCCGCTTGTGGTTGACAGGACCCCTATGCTCTGCAAGGGTACTACAT 1725
1621 ACTATCATCATCCGCTTGTGGTTGACAGGACCCCTATGCTCTGCAAGGGTACTACAT 1680
1726 GAAGTGAGCCAAAGATACGTGTTTATGACGCAAGAGCTAATCTTGTGGGAGCTT 1785
1681 GAAGTGAGCCAAAGATACGTGTTTATGACGCAAGAGCTAATCTTGTGGGAGCTT 1740
1786 GAAATTTAGAGGCGAGTAGGCTTGGAAACCTCTAGGGTTACTTCTTATATACGA 1845
1741 GAAATTTAGAGGCGAGTAGGCTTGGAAACCTCTAGGGTTACTTCTTATATACGA 1800
1846 GGTTCAGTGAAGAACGCTATCTCACTGCTTGGGAAAGAAAGAAAGAGCTTTGAA 1905
1801 GGTTCAGTGAAGAACGCTATCTCACTGCTTGGGAAAGAAAGAAAGAGCTTTGAA 1860
1906 AAACCTATAAGGAAAAACCAAGCATGTTGTCCTGAAAGAAAGAAAGAGAGATGAA 1965
1861 AAACCTATAAGGAAAAACCAAGCATGTTGTCCTGAAAGAAAGAAAGAGAGATGAA 1920
1966 ACAAACTTAGACCTAGTAGAGGACAGCATGTCAGATGTTTCCATGACACTCGGAA 2025
1921 ACAAACTTAGACCTAGTAGAGGACAGCATGTCAGATGTTTCCATGACACTCGGAA 1980
2026 GCCGCTGCGCAGAGACAGAAATGTTTACAGCAAGCAAGCATGTTGTCATGCGTGAATTT 2085
|||||

```

Db 1981 GCCGGTGGCCAGGAAGAATGATACAGCAAGATAGTTGTGATATGCGTAATTT 2040
Qy 2086 CGAAGTACGCTTCATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2145
Db 2041 CGAAGTACGCTTCATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2100
Qy 2146 GAGGTGGAGATTTACATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2205
Db 2101 GAGGTGGAGATTTACATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2160
Qy 2206 TTAATCGGCTTTAAATACGGCCCTCTACAGCCAGTGCATCTCCCGCTAC 2265
Db 2161 TTAATCGGCTTTAAATACGGCCCTCTACAGCCAGTGCATCTCCCGCTAC 2220
Qy 2266 TACAAAGGTCCTCGCTCTGATGAGTTGACCTTACAGCCCTTCTCTCACTTCC 2325
Db 2221 TACAAAGGTCCTCGCTCTGATGAGTTGACCTTACAGCCCTTCTCTCACTTCC 2280
Qy 2326 CGAAGTACGCTTCATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2385
Db 2281 CGAAGTACGCTTCATCTCTGATCCATCGTCGGGGCATTTGACATTTGAACCCGTACTTTA 2340
Qy 2386 ACACTTACCTTCCCAAGACTACGATTTCTGTCGCCCTCTCTCTCATGACAGCGGAG 2445
Db 2341 ACACTTACCTTCCCAAGACTACGATTTCTGTCGCCCTCTCTCTCATGACAGCGGAG 2400
Qy 2446 TTGTTTGAAGAGCTGGAACAAAGCAAGCAGCCTGATGCGGAGACAGCACTGGCATT 2505
Db 2401 TTGTTTGAAGAGCTGGAACAAAGCAAGCAGCCTGATGCGGAGACAGCACTGGCATT 2460
Qy 2506 ACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCTGTCGCCCAAGACTTC 2565
Db 2461 ACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCTGTCGCCCAAGACTTC 2520
Qy 2566 TTGTTTAAATATGCCAGGGGGAATGCCAAATGCCGCTCTGATGACACAGCTTAAG 2625
Db 2521 TTGTTTAAATATGCCAGGGGGAATGCCAAATGCCGCTCTGATGACACAGCTTAAG 2580
Qy 2626 AACATCGCAGATTTAGCAGCCTCTGACAAAGCAGCTCAGATTTCTGGGAAATGCT 2685
Db 2581 AACATCGCAGATTTAGCAGCCTCTGACAAAGCAGCTCAGATTTCTGGGAAATGCT 2640
Qy 2686 GCAATATCCCAACAGCTTTATGATTTCAATTCACACCTCTTTTGAGAGAGTGTATCAAA 2745
Db 2641 GCAATATCCCAACAGCTTTATGATTTCAATTCACACCTCTTTTGAGAGAGTGTATCAAA 2700
Qy 2746 GGAAGAGGGAAGAGTGA 2763
Db 2701 GGAAGAGGGAAGAGTGA 2718

RESULT 6
AB017635 2858 bp mRNA linear ROD 08-JAN-2000
LOCUS AB017635 Cricetus griseus mRNA for ERCC4, complete cds.
DEFINITION AB017635
ACCESSION AB017635
VERSION AB017635.1 GI:6683475
KEYWORDS ERCC4.
SOURCE Cricetus griseus CHO cell_line:CHO-9 cDNA to mRNA.
ORGANISM Cricetus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetus.
REFERENCE 1 (bases 1 to 2858)
AUTHORS Hayashi,T.
TITLES CHO ERCC4 cDNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 2858)
AUTHORS Hayashi,T.
TITLES Direct Submission
JOURNAL Submitted (11-SEP-1998) Tsuyuko Hayashi, Institute of Development,
Aging and Cancer Tohoku University, Molecular Genetics, 4-1
Seiry-machi, Aoba-ku, Sendai, Miyagi 980-77, Japan

```

```

FEATURES
Source
Location/Qualifiers
1..2858
/organism="Cricetus griseus"
/db_xref="taxon:10029"
/cell_line="CHO-9"
/cell_type="CHO"
1..2858
/gene="ERCC4"
43..2784
/gene="ERCC4"
/codon_start=1
/product="ERCC4"
/db_protein_id="BAA89229.1"
/db_xref="GI:6683475"
/translation="MDGISAIVKRAMAPLLEVEROLVLELSDGLVVCARGLADRL
LHFLRLHCHPACLVLVNTPAEVEEFINOLKEVEHLPRVYNEITNSREYVT
OGGIFATSRILVDFLDRIPDLITGLIVYRHRITSCOEAFILRLPRONKRGF
IKAFIDNAVAFDFGCHVERVMNLFVRKLYLMKRFVAVNSFLFOHKEPVEIHWSM
TPAMISTOTAILDIINACLEIKCHNSLEVEDLSFNALGKPDRTIRHYLDPMHO
LGAKTSVLDDILIRLLOYSQIDCVTFNLNLESFRATEKYPGONGSLFLDAST
MEVNARARVVPDVINKRAKSESSEGEETKELEVESNPMWELSEVLKETIEN
KESBALGEPQVILICASDRTCCOLRDYILAGAVFLRLRYRTEKDSABEYVNL
RKSDGPRTKSKDRKPRKDTNKRERASRKAPRKRELITQVMGAEKPEEAGAE
EDDOORATSSPEGGEIHEAFDNLNLSDSAYGLKEPLTIHPLVGCSDPYALRV
LHEVEPVLVLYDAELTFVROLLEYRASRPGKRLRYVFLYGGSTERORLYLRKEK
EAEKELIREKASVYPERGREDTDLIARGLVSDAPADTKRAGGEOHNGPSTV
VDHREPSLESLIHRGIDIEVYTLVEGYIILPELCVERKYSDLTGLNSGRITS
OCLAMSRVYRPPVLLIEFDGKPEPSLAPRSFQESSVSKLTLTLHPRLLRL
WCPSPHATLAEFLERLONKPOPDAAATAMATLTAOSETLPESDKYNPGDPVLMKPGIN
AKNCHSINMNIHVKNIAELASLOSIPDSILGHAGNANVOLYDFLHTVAYDVYSGRVKR"
BASE COUNT 759 a 722 c 776 g 601 t
ORIGIN
Query Match 50.1%; Score 1932.2; DB 10; Length 2858;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 2253; Conservative 0; Mismatches 503; Indels 9; Gaps 1;
Qy 1 GGAAGAGCTTCATGAGTGCAGGGCAGCCGCTGAGCGATTTGCCATGCGCGCTGCTG 60
Db 31 GGAAGAGCTTCATGAGTGCAGGGCAGCTATGCGCTTGCAGAAAGTGGCCCACTGCTG 98
Qy 61 GAGTACGAGCAGACAGCTGCTGGAACCTGCTGACACTGACGGCTAGTAGTGTGCGCC 120
Db 91 GAGTACGAGGAGCAGCTGCTGGAACCTGCTGACAGCGAGCGGCTAGTGTGTGCGCC 150
Qy 121 CGCGGGCTCGGCGGAGCGGCTCTCTTACACCTTTCTCCAGCTGCACCTGCCACCCAGCC 180
Db 151 CGCGGGCTTGGTGCAGACCGGCTCTCTTACCATTTCTGCGGGCTGCACCTGCCACCCGCGC 210
Qy 181 TGCTGTGCTGCTGCTGCACACAGCAGCGGCGGAGAGAGATTTTATTAATACAGCTG 240
Db 211 TGCTGTGCTGCTGCTGCACACAGCAGCGGCTGAGGAGATTTTATTAATACAGCTG 270
Qy 241 AAGATAGAAGAGTGAACACCTCTCGCGGTGTAAACAATTAAGAACAGCAAGCAAGT 300
Db 271 AAGATAGAAGAGTGAACACCTCTCTCCGAGCTGTGACAAATTAAGAACAGCAAGT 330
Qy 301 CGCTATGAAGTTTACACACAAGAGTGTATTAATTTGCGACAGTACAGTACTTGTGTT 360
Db 331 CGCTATGAAGTGTATACACAAGAGTGTATTAATTTGCGACAGTACAGTACTTGTGTT 390
Qy 361 GACTTCTTGAGTGTATTAATACCTTCAAGATTTATTTACTGGCATCTTGTTATAGAGCC 420
Db 391 GATTTCCTTGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 450
Qy 421 CACAGATTAATCGAGTGTGTCAGAGAGCATTAATCTTGCGCCCTTTGCGCAGAAAAC 480
Db 451 CACAGATTAATCGAGTGTGTCAGAGAGCATTAATCTTGCGCCCTTTGCGCAGAAAAC 510
Qy 481 AAGCGTGTATTAAGCTTTTACAGACATCTGTGCTTTGATCTGATGATTTTGT 540

```

Db 511 AAGCGTGCCTTCATCAAAAGCTTTCACAGACAAAGCTGTGGCTTTTGAGACACTGGTTTTGT 570
Qy 541 CATGTGAAAGAGTGTAGGAATCTTTTGTGAGAAAGCTGTATCTGTGGCCAGGTTTC 600
Db 571 CATGTGAAAGAGTGTATGAGAACTTTTGTGAGAAAGCTGTATCTGTGGCCAGGTTTC 630
Qy 601 CATGTGAGCACTAACTCTATTTTGTAGACAGCAAAAGCTGTAGAGTTGTAGAAATCCATGTT 660
Db 631 CATGTGAGCACTAACTCTGTGTGAGCAACACAGCCGTGAAGTGTAGAGTTGTACAGTG 690
Qy 661 TCTATGACAACTTACCTAGCTGTCTATACAGACTGCTATACCTGAGCAATTTTAAATGATGT 720
Db 691 TCTATGACAACTTACCTAGCTGTCTATACAGACTGCTATACCTGAGCAATTTTAAATGATGT 750
Qy 721 CTAAAGCACTAAATATGCAATATCCATATGCTGTGAAGTGAAGATTTATCTTATAGAAAT 780
Db 751 CTGAAGCACTAAATATGCAATATCCATATGCTGTGAAGTGAAGATTTATCTTATAGAAAT 810
Qy 781 GCTATTTGAAAACTTTTGTAGACAAAGCAATCCGCCATTTATCTGGAATCTTTTGTGGCAGCAG 840
Db 811 GCTTGTGGAAAGCGTTGTGACAAAGCAATCCGCCATTTATCTGGAATCTTTTATGGCAGCAG 870
Qy 841 CTGTGAGCCAAAGCACTAAATCTCTTATGAGATTTGAAGATTTACGAACCTTGTCTGAG 900
Db 871 CTGTGAGCCAAAGCACTAAATCTCTTATGAGATTTGAAGATTTACGAACCTTGTCTGAG 930
Qy 901 TATCTCTCTCATGATATGTGTGACATTTCTTATCTCTGGAATCTCTGAGCAACG 960
Db 931 TATCTCTCTCATGATATGTGTGACATTTCTTATCTCTGGAATCTCTGAGCAACG 990
Qy 961 GAAAAAGCTTTTGTGACAAATTCAGGTTGGCTGTCTTCTGATCCAGCAGCCGATGTGTT 1020
Db 991 GAGAAAGTCTTTGTGACAAATTCAGGTTGGCTGTCTTCTGATCCAGCAGCCGATGTGTT 1050
Qy 1021 ATAAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAATATGATTAAGAAAAAGAAAA 1080
Db 1051 GTGAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAATATGATTAAGAAAAAGAAAA 1110
Qy 1081 ATATCTGAAAAAATGGAATTTAAAGAGGAGGAGAAACAAAAAGAAAGAACTGTCTGAGAA 1140
Db 1111 ATGTCT-----GAGAGTGCAAGAGGAGGAGAAACAAAAAGAAAGAACTGTCTGAGAA 1161
Qy 1141 AGCAACCCAAAGTGGGAGGCTGACTGATTAATAAGAAATTTGAGGAGAAAAATTAAG 1200
Db 1162 AGCAACCCAAAGTGGGAGGCTGATTAATAAGAAATTTGAGGAGAAAAATTAAG 1221
Qy 1201 GAGAGTGAAGCTCTGTGTGTGTCAGAGTCAAGTACTGATTTGTCAGAGTGAACGAGAA 1260
Db 1222 GAGAGTGAAGCTCTGTGTGTGTCAGAGTCAAGTACTGATTTGTCAGAGTGAACGAGACT 1281
Qy 1261 TGTTCCTCAGCTGAGACTATATCACTCTGTGAGCGGAGGCTTCTTATTTAGAGCTCTAC 1320
Db 1282 TGTTCCTCAGCTGAGACTATATCACTCTGTGAGCGGAGGCTTCTTATTTAGAGCTCTAC 1341
Qy 1321 AGGAAAACTTTGAGAGAGATTAAGCAAGCTGAAAGTGTGATGAATTTAGAGAGAA 1380
Db 1342 AGGAAAACTTTGAGAGAGATTAAGCAAGCTGAAAGTGTGATGAATTTAGAGAGAG 1401
Qy 1381 GACACTTTCAAGAGAAATTAAGCAATCTCAAAAGAACCTTAAGACCCCAAAAGAAAGAA 1440
Db 1402 GACACTTTCAAGAGAAATTAAGCAATCTCAAAAGAACCTTAAGACCCCAAAAGAAAG 1461
Qy 1441 CGGCGTTTACCAAGAAAGAAACCTTCAAAAGAAAGAAAGAAAGTGTGACTTAACTCA 1500
Db 1462 CGTGCCTTACCAAGAAAGGAGGCTCTTAAAGAGAAAGCGAGAGCTAATCTGTAGCGCA 1521
Qy 1501 ATGTGAGAAAGCTGAAAGAGCTGAGAGAGAGATGTGAGAGAGAGATGTCTGCA 1560
Db 1522 GTGATGAGAGAGCTGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1581
Qy 1561 GAAATTAAG 1620
Db 1582 CAAGCCACCAAGTATGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1641

Qy 1621 AATTGTATCGAGATGCTGCTTTGCGAATCTGAAAGAAACCCCTCATATCATCCAGCG 1680
Db 1642 AATTGTATCGAGATGCTGCTTTGCGAATCTGAAAGAAACCCCTCATATCATCCAGCG 1701
Qy 1681 CTTCGTGGTTGAGGAGACCCCTATGCTGTGACAAAGGTTACTATCAATGAGGAGCA 1740
Db 1702 CTTCGTGGTTGAGGAGACCCCTATGCTGTGACAAAGGTTACTATCAATGAGGAGCA 1761
Qy 1741 TACGTGTTCTTATGAG 1800
Db 1762 TACGTGTTCTTATGAG 1821
Qy 1801 AGTAGAGCTGAGAAACCTGTGAGGTTTCTTCTTATATGAGAGAGAGAGAGAGAG 1860
Db 1822 AGTAGAGCTGAGAAACCTGTGAGGTTTCTTCTTATATGAGAGAGAGAGAGAGAG 1881
Qy 1861 CAAGCTATCTCATGCTTTTGGGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1882 CAGCGTTATCTGACCGCTCTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1941
Qy 1921 AAGCAAGATGTTGTTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 1942 AAGCAAGATGTTGTTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2001
Qy 1981 GTAAGAGGAGAGAGATCTGAGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTT 2040
Db 2002 GCAAGAGGAGAGAGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTT 2061
Qy 2041 CAGAAATGTTTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2062 CAGAAATGTTTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2121
Qy 2101 TCTGTGATGATGTCGTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2122 TCTGTGATGATGTCGTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2181
Qy 2161 ATCTCTACTCCAGAAATGTCGTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2182 ATCTCTACTCCAGAAATGTCGTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
Qy 2221 AATTAAGGAG 2280
Db 2242 AATTAAGGAG 2301
Qy 2281 CTTCGTGATGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Db 2302 CTTCGTGATGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2361
Qy 2341 CAGGAGATCTCCAG 2400
Db 2362 CAGGAGATCTCCAG 2421
Qy 2401 AACTTAAGGAG 2460
Db 2422 CACTGAGGAG 2481
Qy 2461 AATTAAGGAG 2520
Db 2482 AATTAAGGAG 2541
Qy 2521 ACCCTTCCGAGTCAAGAGAGATTAATCTGTGTTCCCAAGAGATTTCTTTAAAGATGCA 2580
Db 2542 ACCCTTCCGAGTCAAGAGAGATTAATCTGTGTTCCCAAGAGATTTCTTTAAAGATGCA 2601
Qy 2581 GGGGATGATGCAAG 2640
Db 2602 GGGGATGATGCAAG 2661
Qy 2641 GCAGGCTGTCAG 2700
Db 2662 GCAGGCTGTCAG 2721

[illegible]

QY	1	GGAGAGACCTTCACATGAGATCAGGGCAGCCGGCTCTGCAGCGATTGCCATCGCGCCGCTGCTG	60
Db	188	GGAAAGGCGGCCCAATGAGAGCCAGGCGCTGTGGGGGAACGGGAGATATGGCGCCGCTGTGG	247
QY	61	GAGTACGAGCGACAGCTGGTGGTCTGGAACTGCTGCACATGCAGGGGCTAGTACTGTGGCC	120
Db	248	GAGTACGAGAGGCGACAGGATGCTGGAACTGCTGCACAGCAAGGGCTGGTGGTGTGTGCC	307
QY	121	CGCGGGCTCGGCGCGGACCGGCTCTCTACCATCTTCTCCAGCTGCATCGCCACCCAGCC	180
Db	308	CGCGAGACTCGGACACCGACGGGCTCTCTACCATTTCTCGGGCTTCACATCGCCACCCGGCC	367
QY	181	TGCTCGTGGTGGTGGTCTCAACACGCGAGCGGGCCGAGGAGSAGTATTTATCAATAGCTG	240
Db	368	TGCTCGTGGTGGTGGTCTCAACACGCGAGCGGGCCGAGGAGSAGTATTTATCAATAGCTG	427
QY	241	AAGATAGAAAGAGTTGAAACACTCCCTCGCGCGGTGAACAAATGAAATGCACAGCAACGT	300
Db	428	AAGATAGAAAGAGTTGAAACACTCTCCCTCGAGGTGACAAATGAAATGCCAAGTAAACGT	487
QY	301	CGCATGAAAGTTTACACACAAAGGTGGTGTATATTTTGCACACAGTAGATCTTGGTT	360
Db	488	CGCATGAAAGTTTACACAGAGGGGTGGTATATTTTGCACAAAGCGAATACTGTTGGT	547
QY	361	GACTTCTTGACTGATAGAAATACCTTCAGATTAATTTACTGCAATCTTGCTGATAGACC	420
Db	548	GATTTCCTTGACAGGTATAAATACCTTCAGATTAATTTACTGCAATCTTGCTGATAGAGGCA	607
QY	421	CACGAATATCGAGCTGTCTCAAGAAAGCATATCTGCGCTCTTTCCGCGCAGAAAC	480
Db	608	CACGAATCATGTAGTCTCTCCAGAAAGCTTATCTTGCGCTCTTTCCGCGCAGAAAC	667
QY	481	AAACGTGGTTTTATTAAGCTTTACAGACAAATGCTGTGGCTTTGATCTGTTTTGGT	540
Db	668	AAAGGGCGGTTTCAATCAAAAGCTTTACACGACAAAGCTTTGCGTTTGACACTGGTTTTGT	727
QY	541	CATGTGGAAGAGTATGACAAATCTTTTTGTGAGAAACTGATCTGTGGCCAAAGTTTC	600
Db	728	CACGTGGGAAGAGTATGACGGAACCTTTTTGTGAGAACTCTTACATGTGGCCAAAGTTTC	787
QY	601	CATGTAGAGTAAACCTATTTTGAAGACGCAACAAACCTGAAGCTTGAAGAAATCCATGTT	660
Db	788	CATGTAGAGTCAACTATTTCTTGGAAACAACAAACCTGAAGCTTGAAGAAATCCATGTT	847
QY	661	TCTATGACACCTTACATGCTTGGTATACAGACTGCTATACTGGACATTTTAAATGCATGT	720
Db	848	TCCATGACTCCCGCAGTGCCTTGGCCATTTCAAGCGGCATACTAGACATCTTAAATGCATGT	907
QY	721	CTAAAGAACTAAATGCCATTAACCCATGCGTCTGGAAGTGAAGATATATCTTAGAANAAT	780
Db	908	CTGAAGAACTGAANTGTCAACACCCGTCACTTGAAGTGAAGATATGTCACTACATAAACC	967
QY	781	GCTATTGGAAACCTTTTGACAGACAAACCCGCAATTACTGGATCTCTTGGGAGCAG	840
Db	968	GCTCTGGGAAGCCATTGCAAGACAAACCCGCAATTACTGGACCTTTTGGGAGCAG	1027
QY	841	CTTGGAGCCAAAGACTAAATCTTACTTACAGAGATTTAAGATATTTAGCACTTTGCTGCAG	900
Db	1028	CTTGGAGCCAAAGACTAAATCTTGGTGGAGAGATTTAAGATATTTAAGAACCTTACTGCAG	1087
QY	901	TATCTCTCTCAGTATGATTTGTCTACATTTCTTAACTCTTGGAAATCTCTGAGACACG	960
Db	1088	TATCTCTCTCAGTATGATTTGTGTACATTTCTGAAATCTCTGAAATCTCTGAGACACCA	1147
QY	961	GAAAGAGCTTTGGTGCAGATTCAGGTGGCTTTCTTGACCTCCGACCCGATGTTT	1020
Db	1148	GAGAGAGCTTTGGTGCAGATTCAGGTGGCTTTCTTGAGTCTGACCCCTCCATGTTT	1207
QY	1021	ATAAATGCTCGAGCANAGGTTTATCATCTTTCTCCAGATGCCAAATAGTAAAAAAMAAAA	1080
Db	1208	GTGAAATGCTCTGCGCCAGATTTACCGAGTTTCCAGATGTCAAACTGACAAAAAGGCCAAA	1267
QY	1081	ATATCTGAAAAAATGAAATTTAAAGAGGGGAAGAAACAAAAAGCAACTGCTCTAGAA	1140

Db 1268 ACATCAGAAAGACGTCGAGTCAGAGTTCAAGAAACAAAAGAACTGGTTCAGAG 1327
 Oy 1141 AACCAACCAGTGGGAGGCTGACGTGAATTTAAAGAAATTTAGGACAGAAATAG 1200
 Db 1328 ACCAATCCAAATGGGAGCGGTGACGTGACGTGAAAGAAATGCAACAAATAAG 1387
 Oy 1201 GAGAGTGAAGCTTGTGGTCCAGTCAAGTACTGATTTGTGCAAGTATGCCAGAA 1260
 Db 1388 GAGACGAAAGCCCTTGTGGTCCAGGCGGCTGTGATTTGTGCAAGTATGCCAG 1447
 Oy 1261 TGTTCACAGCTGAGAGTATATCACTCTTGGAGCGGAGCCCTTATTTAGAGCTTAC 1320
 Db 1448 TGTCTCCAGCTGAGAGTACTGTCTCCGCGGAGCAAACTCTCTGTGGAGCTTAT 1507
 Oy 1321 AGGAAACCTTTGAGAGATAGCAAAAGCTGAAAGAGTGTGATGAATTTAGGAAGAA 1380
 Db 1508 AGGAAACCTTTGAGAGATAGCAAAAGCTGAAAGAGTGTGATGAATTTAGGAAGAGG 1567
 Oy 1381 GACAGTTCAAGAGATTTAGAAATCTCAAAAGACCTAAAGACCCCAAAACAAAGAA 1440
 Db 1568 GACGCGCCGAGAGAGATCAAAATCTGCAAAAGACCAAAAGCCCTCCAAACAAAGAG 1627
 Oy 1441 CGGCTTCTACCAAGAAAGAAC---CTCAAAAGAAAAAGAGAGTTGACCTTAACT 1497
 Db 1628 CGTCTCCGCGCAAAAGGCGGCCCTCTCAAAAGGAGAAAGCAAGAGCTGACTGTGACG 1687
 Oy 1498 CAATAGTGAAGAAACCTGAAGATGGAAGAGAGAGTGTGAGAGAGATATCT 1557
 Db 1688 CAGTTTGGGGTCCGCTGAGAGCGCGCGGAGCAAGAGCTTTGAGAGAAACCTGTGC 1747
 Oy 1558 CGAGAAATTAAGCAGTAGCCAGAAAGCTGCCGGAAGAAATTAAGCATGAAGATTTGAT 1617
 Db 1748 AGACAGACCAAGTAGACCCAGAAAGCTCGGGGTAGAGATTAAGGTGATCAATTTGAT 1807
 Oy 1618 GTAAATTTGTATCGGATGCTCTTTCGAAATCTGAAAGAACCCCTCACTATCATCAT 1677
 Db 1808 TTAATATGTCTGTGATGCGGCTGTGATGCTGAAAGAACCCCTGACATCATCATCAT 1867
 Oy 1678 CCGCTTGTGGTGGAGGACCCCTATGCTGTGACAAAGGCTCTCATATGAGTGGAGCA 1737
 Db 1868 CCGCTTGTGGTGGAGGACCCCTATGCTGTGACAAAGGCTCTCATATGAGTGGAGCA 1927
 Oy 1738 AGATAGTGTCTTATGAGCAGAGCTAACCTTTGTTGCGAGCTTGAATTTACAG 1797
 Db 1928 AGATAGTGTCTTATGAGCAGAGCTAACCTTTGTTGCGAGCTTGAATTTACAG 1987
 Oy 1798 GCGATAGGCTTGGGAAAGCTGTGAGGTTTACTTTTATATACGAGGTTCACTGAG 1857
 Db 1988 GCTACAGGCTTGGGAAAGCTGTGAGGTTTACTTTTATATACGAGGTTCACTGAG 2047
 Oy 1858 GAACAGCTATCTACTGCTTTGGGAAAGAAAGAAAGCTTTGAAAGAACTCATTAAG 1917
 Db 2048 GAGCAGCCCTGCTACTGCTGTGCGGAAAGAAAGAAAGCTTTGAAAGAACTCATTAAG 2107
 Oy 1918 GAAAAAGCAAGCATGTTGTCCCTGAAGAAAGAGAGAGATGAACAAACTTAGAC 1977
 Db 2108 GAAAAAGCAAGCATGTTGTCCCTGAAGAAAGAGAGATGAACAAACTTAGAC 2167
 Oy 1978 CTAGTAAGAGCAGCATCTGCAAGATTTTCCACTGACACTGCGAAAGCCGCTGAGC 2037
 Db 2168 CTAGTAAGAGCAGCATCTGCAAGATTTTCCACTGACACTGCGAAAGCCGCTGAGC 2227
 Oy 2038 GAACAGATGTTGACAGCAAGATAGTTGATGATGCGGATGCGGATGCGGATGAGCT 2097
 Db 2228 GAGCAGATGTTGACAGCAAGATAGTTGATGATGCGGATGCGGATGCGGATGAGCT 2287
 Oy 2098 CCATCTGATCCATCTGTGGGAGATTTGACATTTGAACCCGCTGACTTTAGAGTTGAGAT 2157
 Db 2288 CCATCTGATCCATCTGTGGGAGATTTGACATTTGAACCCGCTGACTTTGAGAGTTGAGAT 2347
 Oy 2158 TACATCTCACTCCAGAAATGTGCTGAGCGCAAGAGTATCATGATTTAACTGCGCTCT 2217

Db 2348 TACATCTCACTCCAGCAAGCTGTGTGTGAGCGCAAAAGTGTGAGCGCACTCATTTGCTCG 2407
 Oy 2218 TTAATTAACGCCGCTTACAGCCAGTGCATCTCATCTCCGCTACTTACAGCTGCC 2277
 Db 2408 CTGACAGTGGCGGCTGTACAGCCAGTGCCTAGCCATGTCTGCTACTTACCGGAGGCT 2467
 Oy 2278 GTGCTTGTAGTGTGATTAACCTTACAGCCCTTCTCTCACTTCCGAGGTGCTTG 2337
 Db 2468 GTGCTTGTAGTGTGATTAACCTTACAGCCCTTCTCTCACTTCCGAGGTGCTTG 2527
 Oy 2338 TTTACAGAGATCTCCAGCAATGATTAATTTCCAAACTCTCTTCACTTCACTTC 2397
 Db 2528 TTTACAGAGATGAGCAGCAGCATGTGAGCTCCAACTTACGCTTCCACCTTGACCTTC 2587
 Oy 2398 CCCAGACTAGCATTTCTGTGTGCTCCCTCTCTCATGCAAGCGGAGTTGTTGAGAG 2457
 Db 2588 CCGCTTGTGCTGTGTGCTGTGCTGCTCCCTCACCCAGCCAGCCAGCTGTTGAGAG 2647
 Oy 2458 CTGAAACAAAGAGAGCAGCTGATGCGGAGCAGCAGCTGCGCATTTACAGAGATCT 2517
 Db 2648 CTGAAACAAAGAGAGCAGCTGATGCGGAGCAGCAGCTGCGCATTTACCGAGACTCA 2707
 Oy 2518 GAAACCTTCCGAGTCAAGAGATTAATCTGTGCTCCCAAGACTTCTGTTAAATG 2577
 Db 2708 GAGACACTCTGAGTCAAGATTAATCTGTGCTCCCAAGACTTCTGTTAAATG 2767
 Oy 2578 CCAGGCTGAATCCAAACCTGCGCTCTTGTATGACACCACTTAAACATTCGAGAA 2637
 Db 2768 CCGGCTGAATCCAAACCTGCGCTCTTGTATGACACCACTTAAACATTCGAGAA 2827
 Oy 2638 TTAGCAGCCCTGTCACAAAGAGCTCAGAGTATTTCTGGGAATTTCTCAAAAGCAAA 2697
 Db 2828 CTGGCAGCCCTGTCACAAAGAGCTCAGAGTATTTCTGGGAATTTCTCAAAAGCAAA 2887
 Oy 2698 CAGCTTTATGATTTTCACTTACACCTCTTTTGAAGAGTGTATCAAAAGGAGGAAA 2757
 Db 2888 CAGCTTCAAGACTCTGCACTGCTATGCAAGTGTGTGTCMAAGCAGAGTGA 2947
 Oy 2758 AAGTGA 2763
 Db 2948 AAGTGA 2953

RESULT 8
 BC026792
 LOCUS
 DEFINITION
 Mus musculus, excision repair cross-complementing rodent repair deficiency, complementation group 4, clone MGC:30522 IMAGE:4911382,
 mRNA, complete cds.
 ACCESSION
 BC026792.1 GI:20073214
 VERSION
 BC026792.1
 KEYWORDS
 MGC.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 3235)
 AUTHORS
 Strausberg, R.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (04-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/INL at: <http://image.llnl.gov>
 Series: IRK Plate: 41 Row: j Column: 17.

FEATURES

Location/Qualifiers
 1..3235

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:30522 IMAGE:4911382"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_lib="MCL CGAP_Sc2"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 24..2744

CDS

/codon_start=1
 /product="excision repair cross-complementing rodent
 repair deficiency, complementation group 4"
 /protein_id="AH26792.1"
 /db_xref="GI:20073215"
 /db_xref="locusid:50505"
 /translation="MAPLEVEROVELLDSDGLVVCARGLGTDLRLHFLRHCHP
 ACVLVINTOPAEFEYFINDOKIEGVHILPRVTNLASREYVTOGIIPTATRI
 LYVDFITGRIPSLITGLIVYRAHRIIESQEFILRLRQKRRFIRATPDNAAF
 DTGFCHVERMRLVRLKLYMPRHVAVNSLEQKREPEVLIHVMATMLIOTAI
 LDILNACLEKLNPSLEVEDLENAAGKPEDKIRHLDLHQLQALSTLSDYD
 LKILRLTLYLSQDVTPLNLESLRATKEVKGQSGMLFDLASTMEFNARAYR
 LPVRLINKKAKTSEKTSPEVCEKKELEKVEALDVLKELEAEKSEALGSG
 PGNVLICASDDRTCOLRDLISAGETFLIRLYRKFEKQKKEEYVNVNREKDKR
 TTKSDRRPRAPKBERASAKRQPLRKQKQELTLYVLSAEPRPDKALEBLCQRT
 SSSPECGVEIKESFDLVNSDAIGILKEPLTTHPLDGSPTALRLVLEVER
 YVLYDAELFVNAOLIYRASRPKPLRYFLIYGGSTEQRYLTLRKEKAEFLI
 REKASVNVVEERDETNLDLARGSLADAPDTRKAGQNGTQSSIVDMREER
 SELPSLIHRGIDIEPVTLVEGDYILTPELCYRKSVSLIGLSHGRVLSCLASMR
 YRRPVLLIEFDPSKPSLAPRGAFQEMSSSDVSXKLTLHPPRLRLMGPSHA
 TAELELKONKPODPAATAMATTADSELTPESDRYNPGDQVILKMPVNAKNSL
 MNOVKNIATLSTLERLTITLHSGSNAKQLHDFLTATVADLVSKRVRK"

BASE COUNT 829 a 838 c 859 g 709 t
 ORIGIN

Query Match 48.5%; Score 1872; DB 10; Length 3235;
 Best Local Similarity 80.4%; Pred. No. 0;
 Matches 2206; Conservative 0; Mismatches 535; Indels 3; Gaps 1;

23 GGCAGCCGGCTCGACGAGATTCATGCGCGCTGCTGAGTACGACGACGAGTGGC 82
 1 GCGTGTGGGCGGAGCGCGGAGTATGCGCGCTGTGAGTACGAGGACGAGGTGC 60
 83 TGGAACTGCTCGACGAGGAGGCTAGTAGTGTGCGCGCGGCTGGCGGCGGACGGC 142
 61 TGGAACTGCTCGACGAGGAGGCTGTGTGTGTGCGCGGAGCTGGCGACCGACGGC 120
 143 TCCCTCAACATTTCTCAGCTGACGACGACGAGCTGCTGCTGTGCTGCTACAA 202
 121 TCCCTCAACATTTCTCAGCTGACGACGACGAGCTGCTGCTGTGCTGCTACAA 180
 203 CGCAGCCGGCGGAGGAGGAGTATTTATCAATCAGCTGAGAGATGAAAGAGTGAACACC 262
 181 CGCAGCCGGCGGAGGAGGAGTATTTATCAATCAGTGAAGATGAAAGAGTGAACATC 240
 263 TCCCTGCGCGGTAAAGTAAATGCAAGCAACAGTGGTATGAAGTTTACACAGAAG 322
 241 TCCCTGCAAGGTGCAAAATGAATGCAAGTACAGTCCGTATGAAGTCTACACAGAAG 300
 323 GTGGTGTATATTTGCGACAGTACTTGTGTGACTTCTGTACTGATGATGAATAC 382
 301 GTGGTATATATTTGCAACAAGCGAATACGTGTGTGATTTCTTGACAGGTGATAC 360
 383 CTTCAGATTTAATTTACTGGCATCTTGGTGTATVAGAGCCACAGAAATATGAGTCTTGTGC 442

361 CTTGAGATTTAATTAATGAGATTCCTCTGTATACAGGACACAGAAATCATGAGTCTGCC 420
 443 AAGAAGCATTCATTTCTGCGCTCTTTGGCCAGAAAACAAACGCGGTATTTATTAACCTT 502
 421 AGGAAGCTTCACTCTCGCGCTCTTCCGCAAGAAAACAGCCCGCTTATCAATCAACCTT 480
 503 TCACAGCAATGCTGTGCGCTTGTATCTGATGCTGTTTGTGATGTCAGAAAGATGTAGAA 562
 481 TCACCGCAACGCTGTGCGCTTGTATCTGATGCTGTTTGTGATGTCAGAAAGATGTAGAA 540
 563 ATCTTTTGTGAGAAACCTGATCTGTGCGCAAGGTTTCATGTAGCAGTAAACTCATTTT 622
 541 ACCTTTTGTGAGAAACCTGATCTGTGCGCAAGGTTTCATGTAGCAGTAAACTCATTTT 600
 623 TAGAAGCAACAAACCTGAAAGTGTAAATTCATGTTTGTATGACACCTACCTGCTGTG 682
 601 TGGAAACAACAAACCTGAAAGTGTAGAGATTCACGTCACATGACATCCGCGCATGCTTG 660
 683 CTATACAGACTGCTATCTGACATTTTAAATGATGCTTAAGCACTAAAGTAAATGCCATA 742
 661 CCAATTCAGAGCGCATCTAGACATCTTAAATGATGCTTGAAGCACTGAATGTACACA 720
 743 ACCCATGCTTGAAGTGAAGATTTATCTTTAGAAAATGCTATTTGCAAAACCTTTTGACA 802
 721 ACCGTCACCTGGAAGTGAAGATTTGTACATAGAAAACGCTTGGCAAGCATTTGACA 780
 803 AGCAATCCGCGCATTTATCTGATCTTTGTGACACAGCTTGAGGCCAAGCACTAAATCCT 862
 781 AGCAATCCGCGCATTTATCTGATCTTTGTGACATGCTTGAGGCCAAGCACTAAATCCT 840
 863 TAGTTCAGATTTAATTAATTTAGCACTTTGCTGCAAGTATCTGTCAGATGATTTGTG 922
 841 TGGTGCAGATTTAATTAATTTAGCACTTTGCTGCAAGTATCTGTCAGATGATTTGTG 900
 923 TCACATTTCTTAATCTTCTGGAATCTGTGAGAGCAACGGAAGAAACCTTTGCTCAGATT 982
 901 TTACATTTCTGAATCTTCTGGAATCTGTGAGAGCAACGGAAGAAACCTTTGCTCAGATT 960
 983 CAGGTGGCTGTTCTGTGACCTCCGACACCTGATGTTTAAATGTCGAGCAAGGGTTT 1042
 961 CAGGTGGCTGTTCTGTGATGCTGACACCTGATGTTTGAATGTCGAGCAAGGGTTT 1020
 1043 ATCATCTTCCAGATGCCAAATGATGATAAAGAAAATATCTGAAAATATGGAATTTA 1102
 1021 ACCGATTCAGATGTCAACTGAACTGAACAAAAGCCAAACATCAAAAGAGCTGAGATTC 1080
 1103 AAGAAGGGAAGAAACAAAGAAAGAACTGCTCTAGAAAGCAACCCAAAGTGGAGAGCAC 1162
 1081 CAGAAGTTCAAGAAACAAAGAAAGAACTGCTCTAGAAAGCAACCCAAAGTGGAGAGCAC 1140
 1163 TGACTGAAGTATTAAGAAATTTAGGAGCAAAATTAAGAGAGTGAAGCTTTGGTGGTC 1222
 1141 TGACTGAAGTCTTAAGAAATTTAGGAGCAAAATTAAGAGAGTGAAGCTTTGGTGGTC 1200
 1223 CAGGTCAAGTACTGATTTGTGCAAGTATGACGCAAGATGTTCCCACTGAGACATATA 1282
 1201 CAGGTCAAGTACTGATTTGTGCAAGTATGACGCAAGATGTTCCCACTGAGACATATA 1260
 1283 TCACCTTTGAGAGCGAGGCTTCTTATTTAGAGCTCTACAGGAACCTTTGAGAGGATA 1342
 1261 TGTCCGCGGAGCAAGAACTTCTGCTGCACTTTTATGGAAGAACTTTTATGAGAGGATG 1320
 1343 GCAAGGCTGAAGAGTGTGATGTAATTTAGGAAGAAAGCACTTTCAAAAGAAATTTAGGA 1402
 1321 GCAAGGCTGAAGAGTGTGATGTAATTTAGGAAGAAAGCACTTTCAAAAGAAATTTAGGA 1380
 1403 AATCTCAAAAGAACTTAAGAGACCCCAAAAGAAAGAGGCTTTTACCAAAAGAAAGAA 1462
 1381 AATCTCAAAAGAACTTAAGAGACCCCAAAAGAAAGAGGCTTTTACCAAAAGAGGAGG 1440
 1463 CC---CTCAAAAGAAAGAAAGAGGCTTTTACCAAAAGAGGCTTTTACCAAAAGAGGAG 1519

D	b	1441	CCCCCTCAAAAGAAAGAAAGCAAGACGTGACCTCTGACGCAAGTTTGTGGGGTCCGCTAGG	1500
Q	y	1520	AACTGGAAAGGAGAGAGATGTGAGGAAGATATCTGTCAGAAATTAACAGTAAAGCCAG	1579
D	b	1501	AGCCGCGCGGAGCAAGAGGCTTTGGAGGAAGACCTGTGCAGACAGACCAGTGTAGCCAG	1560
Q	y	1580	AAAGCTCCCGGAAAGAAATTAAAGCATGAAGAAATTGATGTAATTGTGATCGGATGCTG	1639
D	b	1561	AAGGCTTCGGGGGTAGAGATTTAAGCTGGAATCAATTGATTTAAATGTGTCTGATGGGG	1620
Q	y	1640	CTTTCGGAATCCTGAAAGAAACCCCTCATTATATCATCCGCTTCGGGGTTCAGCAGACC	1699
D	b	1621	CCATATGCGCATCTGAAAGAGCCCTGATCATTGATCATCCGCTTCAGGTTGACAGCCAGC	1680
Q	y	1700	CCATATGCTTGACAAAGGATCTACATGAAGTGGAGCCAAAGATAGCTGGTTCTTTATGACG	1759
D	b	1681	CCATATGCCCCGTAGAGCGGGTACTCTCAGAGGTGGAGCCCAAGATACGTGGTTCTGATATG	1740
Q	y	1760	CAGAGCTAACCTTTGTTCCGACGCTTGAAATTTACAGGGCCGGTGTGGCTGGGAACCTC	1819
D	b	1741	CAGAGCTCACCTTTGTTCCGCACTGGAAATCTACAGGGCTTAGAGGCTTGGGAAACCTC	1800
Q	y	1820	TGAAGGTTTACTTTCTTATATACGAGAGTTTCAACTGAGGAACAAAGCATATCTCACTGGT	1879
D	b	1801	TGAGGGTTTACTTCTTATATACGAGAGGCTCCACGGAAAGAGAGCGTCACTGACTGCTC	1860
Q	y	1880	TGCGGAAAGAAAGAGAGCTTTGAAAAAATCTAATAGGGAAAAAGCAAGATGGTTGTC	1939
D	b	1861	TGCGGAAAGAAAGAGAGAGCTTTGAAAAAATCTAATAGGAAAAAAGCTAGCTGGTTGTC	1920
Q	y	1940	CTGAGAGAAAGAGAGAGAGATGAAGAAACAATTTAGACTAGTATAGAGGCACAGATCTG	1999
D	b	1921	CCGAGAGAGCGGAGAGGACGAGATGAACCAACCTGTGACTTACCAAGAGGCTACAGAGCC	1980
Q	y	2000	CAGATGTTTCCATCTGACACTCGGAAAGCCGGTGGCAGAGAACAGAAATGTACACAGCAA	2059
D	b	1981	TGAGTCTGCCAGGAGACACTCGSAAAGCCGGTGGCCAGAGACAGAAATGGACCACTGCA	2040
Q	y	2060	GCATAGTTGTGATATGCTGAAATTTGGAAGTGAAGCTTCATCTCTGATTCATGCTGGG	2119
D	b	2041	GCATCGGTGGAGACATAGCCGAGATTTGGAGAGAGACTCCATCTCTGATTCATGCGGG	2100
Q	y	2120	GCATTTGACATTTGAACCCGTGACTTTAGAGTTGGAGATTTACATCTCTCACTCAGAAATGT	2179
D	b	2101	GTAATTGACATAGAGCCAGGTGACTGTGAGGTGGGTGATTTACATCTGTGACACTGAGCTGT	2160
Q	y	2180	GGGTGGAGGCCAAGAGATACAGATTAATGAGGCTTTAAATAGGCCGCCCTCTACA	2239
D	b	2161	GTGTGGAGGCCAAAGATGTGAGGAGCTTAATGTGCTGCTGCACAGTGGCGGCTGTACA	2220
Q	y	2240	GCCAGTGCATCTCATATGTCGCCGTACTACAAAGCGTCCGCTGCTTGTGATTAAGTTGACC	2299
D	b	2221	GCCAGTGCCTAAGCATATGTCTCGGTACTTACCGGAGGGCTGTGCTGCTAATGTAGTTGACC	2280
Q	y	2300	CTAGCAAGCCTTTCTCTCTCACTTCCGAGGTGCTTGTGTTAGAGAAATCTCCAGCAATG	2359
D	b	2281	CGACCAAGCCCTTCTCTCTGCGCCGACGAGGTGCTCTTCCAGGAGATGAGCAGACGCG	2340
Q	y	2360	ACATTAATTCCAAACATCACTCTTACACTTCACTTCCCGAGACTACAGGAATCTGTGCT	2419
D	b	2341	ATGTGAGCTCCAAAGCTTACAGCTCTCTCACTTGCATCTCCGCGTGTGGGCTGTGCTGT	2400
Q	y	2420	GCCCTCTCTCAATGCACAGCGCGAGATTTGTTGAGAGCTGAACCAAGCAAGCCACAGC	2479
D	b	2401	GCCCTCAACCCACAGCCACAGCGGAGCTGTTTGAAGAGCTGAAGCAAGAACACAGC	2460
Q	y	2480	CTGATGCGGCGACAGCACTGTGCCATTACAGCAGAAATCTGAAACCTTCCGAGTCAGAGA	2539
D	b	2461	CTGATGCGGCGACAGCCATGGCCATACACGCAAGACTCAAGACATGCTGATGTCGAGACA	2520
Q	y	2540	AGTAAATCCGTTGGGCCCAAGACTTCTTTTAAAAATGCCAGGGGTGAATGCCAAAAACT	2599
D	b	2521	GATTAATCCCGGGCCCCCAAGACTTTGTGCTGAAGGTGCTTGGGTCAACGCCAAGAACT	2580

QY	2600	GCCTCTCTTATGACACAGCTTAAAGCATTCGAGATTTGACAGCCCTGTCACAAAGC	2659
Db	2581	GTCCGTCCCTGATGAAACCAAGTAGAAGAACATTCGACAGACTGCCACCTGTCCCTGGAAAC	2640
QY	2660	AGCTCAGCAGTAATTCCTGGGGAATGCTGCMAATGCCMAACAGCTTTATCATTTCAATTCACA	2719
Db	2641	GACTCACAACCACTCCCTGGGACATCGGGGAATATGCCAACACACTGCACACTTCTCTGCACA	2700
QY	2720	CCTCTTTTGCAAGAGCTGATCAAAAGGAAGAAAGGMAAAGTGA	2763
Db	2701	CTGCCTATGACAGACTGGTGTCCAAAGGACAGATGAGAAAGTGA	2744
RESULT 9			
LOCUS	AX332256	35641 bp	DNA
DEFINITION	Sequence 2765 from Patent WO0194629.	Linear	PAT 09-JAN-2002
ACCESSION	AX332256		
VERSION	AX332256.1	GI:18122890	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mumtaz, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrijan, S., Soppet, D.R. and Weaver, Z.		
TITLE	Cancer gene determination and therapeutic screening using signature gene sets		
JOURNAL	Patent: WO 0194629-A 2765 13-DEC-2001;		
FEATURES	Avalon Pharmaceuticals (US)		
SOURCE	Location/Qualifiers		
	1..35641		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	10575 a 7169 c 7323 g 10574 t		
ORIGIN			
	Query Match 47.4%: Score 1829.4; DB 6; Length 35641;		
	Best Local Similarity 99.4%: Pred. No. 0;		
	Matches 1836; Conservative 0; Mismatches 11; Indels 0; Gaps 0;		
QY	2011	ACTGACACTCGGAAGCCGCTGGCCAGAGACAGATGTGTACACAGCAAGCATAGTTGTG	2070
Db	29522	ACTTACTTTTCTCTGTAGTGTGGCCAGGACAGATGTGTACACAGCAAGCATAGTTGTG	29581
QY	2071	GATATGCGTGAATTTTCGAAGTGACGCTTCATCTGTATCATCTCGTGGGCAATTGACATT	2130
Db	29582	GATATGCGTGAATTTTCGAAGTGACGCTTCATCTGTATCATCTCGTGGGCAATTGACATT	29641
QY	2131	GAACCCGAGACTTGTAGAGTTGGAGATATCAATCCATCACTCCAGAAATGGGTGGAGGCG	2190
Db	29642	GAACCCGAGACTTGTAGAGTTGGAGATATCAATCCATCACTCCAGAAATGGGTGGAGGCG	29701
QY	2191	AAGACTATCAGTATTTAATGAGGCTCTTAAATAAAGCGCCCTCTACAGCCAGTGCATC	2250
Db	29702	AAGACTATCAGTATTTAATGAGGCTCTTAAATAAAGCGCCCTCTACAGCCAGTGCATC	29761
QY	2251	TCCATGTCGCCGCTACTACAAAGCGTCCGCTCTGTGATTAAGTTGACCTTACGAAGCCT	2310
Db	29762	TCCATGTCGCCGCTACTACAAAGCGTCCGCTCTGTGATTAAGTTGACCTTACGAAGCCT	29821
QY	2311	TTCTCTGTCACGTCGCCGAGGTGCTTTGTTAAGAGGATCTCCACCAATGACATTAGTTCC	2370
Db	29822	TTCTCTGTCACGTCGCCGAGGTGCTTTGTTAAGAGGATCTCCACCAATGACATTAGTTCC	29881
QY	2371	AAACTCACTCTTCTTACACTTCCCTCCAGACATACAGATTTCTGTGTGCCCTTCTCT	2430
Db	29882	AAACTCACTCTTCTTACACTTCCCTCCAGACATACAGATTTCTGTGTGCCCTTCTCT	29941
QY	2431	CATGCAAGCGGAGATTGTTTGAAGAGCTGAACCAAGCAAGCCACAGCTGATGCGCG	2490

Db 29942 CATGCAACGGCGAGTGTGTTGAGAGCTGAACAACGACCAACGCTGATGGGCG 30001
QY 2491 ACAGACAGTGGCATTAACACAGATTCGAAACCTCCGAGTGCAGAGATATATCT 2550
Db 30002 ACAGCAGTGGCATTAACACAGATTCGAAACCTCCGAGTGCAGAGATATATCT 30061
QY 2551 GGTCCCAAGACCTTGTGTTAAATAATGCGAGGGGTGAATGCCAAAACGCGCTCTTG 2610
Db 30062 GGTCCCAAGACCTTGTGTTAAATAATGCGAGGGGTGAATGCCAAAACGCGCTCTTG 30121
QY 2611 ATGCACACAGTGAAGACATCGCAGATTTAGCAGCCCTGTCAAGACAGAGCTCACAGT 2670
Db 30122 ATGCACACAGTGAAGACATCGCAGATTTAGCAGCCCTGTCAAGACAGAGCTCACAGT 30181
QY 2671 ATTCGGGGAGATGCTGCAAAATGCAACAGCTTTATGATTGATTCACACCTCTTTGCA 2730
Db 30182 ATTCGGGGAGATGCTGCAAAATGCAACAGCTTTATGATTGATTCACACCTCTTTGCA 30241
QY 2731 GAAGTCGTATCAAAAAGGAAAAAGGAAAAAGTGAACAGTGTGCTTTTCTATCCAT 2790
Db 30242 GAAGTCGTATCAAAAAGGAAAAAGGAAAAAGTGAACAGTGTGCTTTTCTATCCAT 30301
QY 2791 GCTGTACTTTTCAGCGGCTCTTCGCAACATCATAGTCAATTAATTAATGTTTG 2850
Db 30302 GCTGTACTTTTCAGCGGCTCTTCGCAACATCATAGTCAATTAATTAATGTTTG 30361
QY 2851 CTATTTTCATCTTTTCCAAATGCTCTTAATGATTTAGCGTGACAGCAGAGATTCG 2910
Db 30362 CTATTTTCATCTTTTCCAAATGCTCTTAATGATTTAGCGTGACAGCAGAGATTCG 30421
QY 2911 TCTGTGAACCTGCAAGTATGACATCACTGAACCTGCTGCTGCTCTTTTCTCTCC 2970
Db 30422 TCTGTGAACCTGCAAGTATGACATCACTGAACCTGCTGCTGCTCTTTTCTCTCC 30481
QY 2971 TGCACCGCTATGCGGGGCTTGCAATGTTCTTTTAAATGAGGTTTGCAGATCAGT 3030
Db 30482 TGCACCGCTATGCGGGGCTTGCAATGTTCTTTTAAATGAGGTTTGCAGATCAGT 30541
QY 3031 AAGTTCCTACAGATGATACAGAAAGTGAAGCTTACCTGATCTTACAGATCTCAT 3090
Db 30542 AAGTTCCTACAGATGATACAGAAAGTGAAGCTTACCTGATCTTACAGATCTCAT 30601
QY 3091 TGAAGAAGATATGCTAGCCTGCGCATGAGCGGTGACAGGAGGAAAAAGACAGCACA 3150
Db 30602 TGAAGAAGATATGCTAGCCTGCGCATGAGCGGTGACAGGAGGAAAAAGACAGCACA 30661
QY 3151 GAAGCTACATTTTAAACAGTCTGTTATCTAGTGCACAAATTAATACAGCTTAAT 3210
Db 30662 GAAGCTACATTTTAAACAGTCTGTTATCTAGTGCACAAATTAATACAGCTTAAT 30721
QY 3211 GCATTTATACCATGTCCTGCTGCTCCAAATCTGCTCTTTCGCTGTTGCTCTGCTGA 3270
Db 30722 GCATTTATACCATGTCCTGCTGCTCCAAATCTGCTCTTTCGCTGTTGCTCTGCTGA 30781
QY 3271 CGCTGAGTGAATGTTGTGTGAAGAAATCATGTTGACCTTTGTCTACAAAGAGCT 3330
Db 30782 CGCTGAGTGAATGTTGTGTGAAGAAATCATGTTGACCTTTGTCTACAAAGAGCT 30841
QY 3331 TGTGGAACCTGAGAAGAAACATCTTTGSCATCTGACAGATTCCTCTACACAT 3390
Db 30842 TGTGGAACCTGAGAAGAAACATCTTTGSCATCTGACAGATTCCTCTACACAT 30901
QY 3391 TTCTTCAGTCAATCTTGTGCTGCTCTAAGGAAATTCATGAGGCTCTCTACT 3450
Db 30902 TTCTTCAGTCAATCTTGTGCTGCTCTAAGGAAATTCATGAGGCTCTCTACT 30961
QY 3451 ACTAATTCAGACAGTCTCTCAAAAACGTTGACTAGTCTTCTAATGACCTTAACATA 3510
Db 30962 ACTAATTCAGACAGTCTCTCAAAAACGTTGACTAGTCTTCTAATGACCTTAACATA 31021
QY 3511 TGTAGCATATCTATTAATTTCAATGTTGCCAAATAGATTTTTTAAAGCAAAATGAATTAC 3570
Db 31022 TGTAGCATATCTATTAATTTCAATGTTGCCAAATAGATTTTTTAAAGCAAAATGAATTAC 31081

QY 3571 CTGTTGCAAAAGTTAATGATGAGAGGCTTAGAATTCGAAATTTGACATATTCA 3630
Db 31082 CTGTTGCAAAAGTTAATGATGAGAGGCTTAGAATTCGAAATTTGACATATTCA 31141
QY 3631 GTCTCTAATATACAGATCCCTAAGTCCAGCTGGCTAGTACAGATTTTTCAGACT 3690
Db 31142 GTCTCTAATATACAGATCCCTAAGTCCAGCTGGCTAGTACAGATTTTTCAGACT 31201
QY 3691 CCTGCTTTTCAGCTCTTTATTCCTTAAGACACAGCATATATCTCTAAGAAATACACC 3750
Db 31202 CCTGCTTTTCAGCTCTTTATTCCTTAAGACACAGCATATATCTCTAAGAAATACACC 31261
QY 3751 TAATTGGCAGTGAAGCAGATGCGACACACTGACCCCTGCGGCGAGAGTGAAGT 3810
Db 31262 TAATTGGCAGTGAAGCAGATGCGACACACTGACCCCTGCGGCGAGAGTGAAGT 31321
QY 3811 TTGCTCTATTACAAAAAGAAAAAGAAATACAACTAAGCTCA 3857
Db 31322 TTGCTCTATTACAAAAAGAAAAAGAAATACAACTAAGCTCA 31368
RESULT 10
AA409578 35641 bp DNA linear PAT 14-JUN-2002
LOCUS AA409578
DEFINITION Sequence 2225 from Patent WO0229103.
ACCESSION AA409578
VERSION AA409578.1 GI:21442283
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
PATENT: WO 0229103-A 2225 11-Apr-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
source location/Qualifiers
1..35641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U75658"
BASE COUNT 10575 a 7169 c 7323 g 10574 t
ORIGIN
Query Match 47.4%; Score 1829.4; DB 6; Length 35641;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 2011 ACTGACACTCGGAAAGCCGCTGGCCAGAAAGATGTACACACCAGATAGTTGTG 2070
Db 29522 ACTTACTTTTCTCTGTAGTGGCCAGAAAGATGTACACACCAGATAGTTGTG 29581
QY 2071 GATATGCGTGAATTTGAAAGTGAAGCTTCATCTGTATCATCTGCGGCAATTGACATT 2130
Db 29582 GATATGCGTGAATTTGAAAGTGAAGCTTCATCTGTATCATCTGCGGCAATTGACATT 29641
QY 2131 GAACCCGCTGATTTAGAGTTGGAATTTACATCTCCTCAGAAATGTCGCTGAGCGCC 2190
Db 29642 GAACCCGCTGATTTAGAGTTGGAATTTACATCTCCTCAGAAATGTCGCTGAGCGCC 29701
QY 2191 AAGAGTACAGATTTAATGCGCTTTAATTAACGGCGGCTCTACAGCCAGTGCATC 2250
Db 29702 AAGAGTACAGATTTAATGCGCTTTAATTAACGGCGGCTCTACAGCCAGTGCATC 29761
QY 2251 TCCATGTCGCGTACTACAAAGCTCCGCTTCTGATTTGACCTAGCAAGCCT 2310
Db 29762 TCCATGTCGCGTACTACAAAGCTCCGCTTCTGATTTGACCTAGCAAGCCT 29821
QY 2311 TTCTCTCTACCTCCGAGGCTCTTGTTCAGAGATCTCCAGCAATGACATTAATGCTC 2370
Db 29822 TTCTCTCTACCTCCGAGGCTCTTGTTCAGAGATCTCCAGCAATGACATTAATGCTC 29881

OY	2371	AACTCAGCTCTTCTTACACTTTCACCTTCCCCAGACTACAGATTCCTGTGGGCCCTCTCT	2430
OY	29882	AAAGTACACTCTTCTTACACTTCTACCTTCCCCACAGTACAGGATTCCTGTGGGCCCTCTCT	29941
OY	2431	CATGCAGCGGGGAGTTGTTTGTAGGAGGCTGAAGAACAGAACCCAGCTCATCGGGC	2490
Db	29942	CATGCAGCGGGGAGTTGTTTGTAGGAGGCTGAAGAACAGAACCCAGCTCATCGGGC	30001
OY	2491	ACAGCACTGGCCATTACAGCAGATTCGTAACCCCTTCCGAGTCAGAGAATTAATCT	2550
Db	30002	ACAGCACTGGCCATTACAGCAGATTCGTAACCCCTTCCGAGTCAGAGAATTAATCT	30061
OY	2551	GGTCCCAAGACTTCTGTGTAAATATGCCAGGGGGGAATGCCAAAACGCGCGTCCTTG	2610
Db	30062	GGTCCCAAGACTTCTGTGTAAATATGCCAGGGGGGAATGCCAAAACGCGCGTCCTTG	30121
OY	2611	ATGCACCACTTTAAGACATGCGAGAATTTAGAGCCCTGTCTCAAGAGAGCTCACAGAT	2670
Db	30122	ATGCACCACTTTAAGACATGCGAGAATTTAGAGCCCTGTCTCAAGAGAGCTCACAGAT	30181
OY	2671	ATTCGGGGGAATGCTGCGAAATGCCCCAAACGCTTATGATTTATTTCACACCTCTTTGCA	2730
Db	30182	ATTCGGGGGAATGCTGCGAAATGCCCCAAACGCTTATGATTTATTTCACACCTCTTTGCA	30241
OY	2731	GAAGCTGATATCAAAAGAAAAAGGAAAAAGTATACAGATCATGGCTGTTTCTTATCCAT	2790
Db	30242	GAAGCTGATATCAAAAGAAAAAGGAAAAAGTATACAGATCATGGCTGTTTCTTATCCAT	30301
OY	2791	GCTCTACTTTTCAGCGGCTCCTTCCAGACATCATAGTCATTTAATTAATTAATGTTTG	2850
Db	30302	GCTCTACTTTTCAGCGGCTCCTTCCAGACATCATAGTCATTTAATTAATTAATGTTTG	30361
OY	2851	CTATTTCATTTCTTTTCCCATGCTCTTAATGATTTGATAGGTGACACAGAACCCAGATTC	2910
Db	30362	CTATTTCATTTCTTTTCCCATGCTCTTAATGATTTGATAGGTGACACAGAACCCAGATTC	30421
OY	2911	TCTCTGAACCTCTGCAGTTAGGATCAGTGAACCTTGCCTGCTCCCTCTTTTCTCTCC	2970
Db	30422	TCTCTGAACCTCTGCAGTTAGGATCAGTGAACCTTGCCTGCTCCCTCTTTTCTCTCC	30481
OY	2971	TGCACCGTCTAAGCCGGGCTTAGCATGTTTCTTTTAAATGAGGTTTGTACAGATCAGT	3030
Db	30482	TGCACCGTCTAAGCCGGGCTTAGCATGTTTCTTTTAAATGAGGTTTGTACAGATCAGT	30541
OY	3031	AAAGTTCCTACAAAGATTTACAGAGAGTGAAGAACTTACCTATCTTAACACATCTCAT	3090
Db	30542	AAAGTTCCTACAAAGATTTACAGAGAGTGAAGAACTTACCTATCTTAACACATCTCAT	30601
OY	3091	TAGAAAGGAATATGCTAAACCTTGGCATGTGACGCGTGCAGGAGGAAAAAGACAGCACAA	3150
Db	30602	TAGAAAGGAATATGCTAAACCTTGGCATGTGACGCGTGCAGGAGGAAAAAGACAGCACAA	30661
OY	3151	GAAGACTACGATTTTAAACAGAGCTCTGTGTTATCTAGGCAACCTAATATACATCTTAAT	3210
Db	30662	GAAGACTACGATTTTAAACAGAGCTCTGTGTTATCTAGGCAACCTAATATACATCTTAAT	30721
OY	3211	GCACCTAATCCATGTCCTGTGGCTCTCCAAATCTGGTCTTGTGCTGTGTCTGTCTGTGA	3270
Db	30722	GCACCTAATCCATGTCCTGTGGCTCTCCAAATCTGGTCTTGTGCTGTGTCTGTCTGTGA	30781
OY	3271	CGCTTGAACCTGATGTTTGTGTAGGAAATCATGTCGACCCCTTGTCTTCAAGAGAGCCT	3330
Db	30782	CGCTTGAACCTGATGTTTGTGTAGGAAATCATGTCGACCCCTTGTCTTCAAGAGAGCCT	30841
OY	3331	TCTGGAACTGAGAGAAACATCTCTTGGCATTCCTGACCAAGTCTCTCTACACAT	3390
Db	30842	TCTGGAACTGAGAGAAACATCTCTTGGCATTCCTGACCAAGTCTCTCTACACAT	30901
OY	3391	TTCCTCAGCTCATCTCTCTGCTGTCTGCTCTTAAGAGAAATTCATGAGGCTTCTCTACT	3450
Db	30902	TTCCTCAGCTCATCTCTCTGCTGTCTGCTCTTAAGAGAAATTCATGAGGCTTCTCTACT	30961

QY	3451	ACTAATTCAGACAGACTCTCCCTCAAAAAGTGGTGTACTAGTCTCTCTAATGACCCCTAACATA	3510
Db	30962	ACTAATTCAGACAGACTCTCCCTCAAAAAGTGGTGTACTAGTCTCTCTAATGACCCCTAACATA	31021
QY	3511	TGTACCATATACATTAATTTCAATTTGTCTCCAAATTTGATTTTAAAGCAAAATGAATTAC	3570
Db	31022	TGTACCATATACATTAATTTCAATTTGTCTCCAAATTTGATTTTAAAGCAAAATGAATTAC	31081
QY	3571	CTGTTTGGCAAAAGTTAATGATGTAAGAGACTCTTAGAATTTCTCAATTTTTCACATATTTCA	3630
Db	31082	CTGTTTGGCAAAAGTTAATGATGTAAGAGACTCTTAGAATTTCTCAATTTTTCACATATTTCA	31141
QY	3631	GTCTCTTAATATACAGAGATCCCTAAGTCAGCTGGCTAGTTACAGAGTTTTTTCAGACTT	3690
Db	31142	GTCTCTTAATATACAGAGATCCCTAAGTCAGCTGGCTAGTTACAGAGTTTTTTCAGACTT	31201
QY	3691	CGTGTCTTCTAGCTCTTAATATCCCTAATACACACACATCATCTCTAGAAATACAAAC	3750
Db	31202	CGTGTCTTCTAGCTCTTAATATCCCTAATACACACACATCATCTCTAGAAATACAAAC	31261
QY	3751	TAAATGGCAGTAGAGCCGAGATGCACACACTGACCCCTGCGGGCGACAGAGTAGACT	3810
Db	31262	TAAATGGCAGTAGAGCCGAGATGCACACACTGACCCCTGCGGGCGACAGAGTAGACT	31321
QY	3811	TTGTCTCTATTACAAAAAGAAAAAGAAATACACCTTAAGCTCA	3857
Db	31322	TTGTCTCTATTACAAAAAGAAAAAGAAATACACCTTAAGCTCA	31368
RESULT 11			
HOMERCC4G		35641 bp	DNA
LOCUS			linear
DEFINITION			PRI 22-MAR-1997
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
exon			
intron			

Db 29702 AAGAGATGATGATTTAATGCGCTCTTAAATAAGCGCGCTCTAGACGACGATC 29761
 QY 2251 TCCATGTCCTGCTACTACAGCGTCCGCTTCTGATTTGAGTTGACCTTAGCAAGCT 2310
 Db 29762 TCCATGTCCTGCTACTACAGCGTCCGCTTCTGATTTGAGTTGACCTTAGCAAGCT 29821
 QY 2311 TCTCTGCTACTTCCGAGGCTGCTTGTTCAGGAGATCTCAGCAATACATTTAGTTC 2370
 Db 29822 TCTCTGCTACTTCCGAGGCTGCTTGTTCAGGAGATCTCAGCAATACATTTAGTTC 29881
 QY 2371 AAACGCTCTTCTTACACTTCCAGACTACGAGATCTGCTGCTGCTCTCT 2430
 Db 29882 AAACGCTCTTCTTACACTTCCAGACTACGAGATCTGCTGCTGCTCTCT 29941
 QY 2431 CATGCAAGCGGAGGAGTGTTCAGAGCTGAACAAGCAAGCAGCTGATGCGGCG 2490
 Db 29942 CATGCAAGCGGAGGAGTGTTCAGAGCTGAACAAGCAAGCAGCTGATGCGGCG 30001
 QY 2491 ACAGCACTGCGCATTAACAGAGATCTGAAACCTTCCGAGTCAAGAGATATATCT 2550
 Db 30002 ACAGCACTGCGCATTAACAGAGATCTGAAACCTTCCGAGTCAAGAGATATATCT 30061
 QY 2551 GGTCCCAAGACTCTTGTAAAAATGCCAGGGGTGAATGCCAAAACTGCCCTCTTG 2610
 Db 30062 GGTCCCAAGACTCTTGTAAAAATGCCAGGGGTGAATGCCAAAACTGCCCTCTTG 30121
 QY 2611 ATGCAACAGCTTAAGAACATCGCAGATTAAGACGCTCTCAAGAGAGCTCACAGCT 2670
 Db 30122 ATGCAACAGCTTAAGAACATCGCAGATTAAGACGCTCTCAAGAGAGCTCACAGCT 30181
 QY 2671 ATTGCGGGAATGCTGCAATGCCAAACGCTTTATGATTTATTCACACCTTTTGA 2730
 Db 30182 ATTGCGGGAATGCTGCAATGCCAAACGCTTTATGATTTATTCACACCTTTTGA 30241
 QY 2731 GAATGCTATCAAAAAGAAAAAGAAAGTGAACAGTGTGCTTTTCTATCCAT 2790
 Db 30242 GAATGCTATCAAAAAGAAAAAGAAAGTGAACAGTGTGCTTTTCTATCCAT 30301
 QY 2791 GCTCTACTTTTACAGGCTCTCTTCCAGACATCATAGTCTATTTATTTGTTG 2850
 Db 30302 GCTCTACTTTTACAGGCTCTCTTCCAGACATCATAGTCTATTTATTTGTTG 30361
 QY 2851 CTATTTCACTTTTCCATAGCTCTTAATGATTTAGCGTGAGCAGAACGAGATTC 2910
 Db 30362 CTATTTCACTTTTCCATAGCTCTTAATGATTTAGCGTGAGCAGAACGAGATTC 30421
 QY 2911 TCTCTGACTCTGCAAGTGAAGCATCACTTGAATGCTGCTGCTCTTTCTCTCC 2970
 Db 30422 TCTCTGACTCTGCAAGTGAAGCATCACTTGAATGCTGCTGCTCTTTCTCTCC 30481
 QY 2971 TGCACGCTATGCGGGGCTTGAAGTCTTTTAAATGAGTTTGCAGATCAGT 3030
 Db 30482 TGCACGCTATGCGGGGCTTGAAGTCTTTTAAATGAGTTTGCAGATCAGT 30541
 QY 3031 AAAGTCTTCAAGATGATTACAGAGTAGAAACTTTACCTGATCTTAACAGATCTCAT 3090
 Db 30542 AAAGTCTTCAAGATGATTACAGAGTAGAAACTTTACCTGATCTTAACAGATCTCAT 30601
 QY 3091 TAGAAGAGAAATGCTAGAGCTGCGATGAGCGGTGAGGAGGAGAAAGAGACACAA 3150
 Db 30602 TAGAAGAGAAATGCTAGAGCTGCGATGAGCGGTGAGGAGGAGAAAGAGACACAA 30661
 QY 3151 GAAAGCTACCATTTTAAACAGCTGCTGTTATCTAGTGAACATAATACAGCTTAAT 3210
 Db 30662 GAAAGCTACCATTTTAAACAGCTGCTGTTATCTAGTGAACATAATACAGCTTAAT 30721
 QY 3211 GCATTTATACCATCTCTGCTGCTCCAAATGCTGCTTGTGCTGCTGCTGGA 3270
 Db 30722 GCATTTATACCATCTCTGCTGCTCCAAATGCTGCTTGTGCTGCTGCTGGA 30781
 QY 3271 CGCTTGAACTGATTTTGTGTGAGAAATCATCTTCTGACCTTTGTCTACAAAGAGCT 3330

Db 30782 CGCTTGAACTGATTTTGTGTGAGAAATCATGTTCTGACCCCTTTGTCTACAAAGAGCT 30841
 QY 3331 TCTGGAACACTGAGAAAGAAATCTCTTTTCCATTCCTGACCACTTCTCTACACAT 3390
 Db 30842 TCTGGAACACTGAGAAAGAAATCTCTTTTCCATTCCTGACCACTTCTCTACACAT 30901
 QY 3391 TCTCTGAGCTCATCTTCTGCTGCTGCTCTTAAGGAATTCATGAGCCCTTCTACT 3450
 Db 30902 TCTCTGAGCTCATCTTCTGCTGCTGCTCTTAAGGAATTCATGAGCCCTTCTACT 30961
 QY 3451 ACTAATTCAGACAGTCTCTCAAAAAGCTGTGAGTACTTCTTAATACCTTAACATA 3510
 Db 30962 ACTAATTCAGACAGTCTCTCAAAAAGCTGTGAGTACTTCTTAATACCTTAACATA 31021
 QY 3511 TGTAGCATATATATATTTCAATTTGTTCCAAATTAATTTTAAAGCAAAATGAATTAC 3570
 Db 31022 TGTAGCATATATATATTTCAATTTGTTCCAAATTAATTTTAAAGCAAAATGAATTAC 31081
 QY 3571 CTGTTTGCAAAAGTTAATGATGAGGAGCTTTAGAAATCTCAATTTTGCATATTTCA 3630
 Db 31082 CTGTTTGCAAAAGTTAATGATGAGGAGCTTTAGAAATCTCAATTTTGCATATTTCA 31141
 QY 3631 GTCTCTAATATCAGAGATCCCTAAGTCCAGCTGCTAGTTACAGATTTTTCAGACT 3690
 Db 31142 GTCTCTAATATCAGAGATCCCTAAGTCCAGCTGCTAGTTACAGATTTTTCAGACT 31201
 QY 3691 CCTGTTTTCAGCTCTTATATCTTAAGACACACAGCATATATCTTGAANAATACAC 3750
 Db 31202 CCTGTTTTCAGCTCTTATATCTTAAGACACACAGCATATATCTTGAANAATACAC 31261
 QY 3751 TAATTGGCAGTGAAGCGGAGATGCAACACTGCAACCCCTGCGGGGAGAGAGTGAGAT 3810
 Db 31262 TAATTGGCAGTGAAGCGGAGATGCAACACTGCAACCCCTGCGGGGAGAGAGTGAGAT 31321
 QY 3811 TTGCTCTATTTACAAAAGAAAAAGAAATACAACTAAGCTCA 3857
 Db 31322 TTGCTCTATTTACAAAAGAAAAAGAAATACAACTAAGCTCA 31368

RESULT 12
 AC010401/c 119172 bp DNA linear PRI 02-NOV-2001
 LOCUS
 DEFINITION Homo sapiens chromosome 16 clone CTD-2135D7, complete sequence.
 AC010401 AC010401.6 GI:16596527
 VERSION
 KEYWORDS HIG.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (02-NOV-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT
 Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 GenBank flat file format but are available as part

Db 86161 TTGCTCTATTACAAAAGAAAAGAAATACAACTAGCTCA 86115

|||||

RESULT 13
AC009173 149849 bp DNA linear PRI 26-APR-2002

DEFINITION Homo sapiens chromosome 16 clone RP11-99H5, complete sequence.
AC009173
VERSION AC009173.8 GI:20330796
KEYWORDS HIC.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
2 (bases 1 to 149849)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 149849)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 149849)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 26 2002 this sequence version replaced gi:19172832.

COMMENT
Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

FEATURES
source Location/Qualifiers
1. 149849
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-99H5"
BASE COUNT 43996 a 31384 c 30828 g 43641 t
ORIGIN

Query Match 47.3%; Score 1826.2; DB 9; Length 149849;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1834; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2011 ACTGACACTCGGAAGCCGGTGGCCAGAGCAAGATGTGTACACAGCAAGCATAGTTGTG 2070
|||||
Db 115652 ACTTATCTTTCTCTGTGTGTGGTGGCCAGGAAGCAATGTGTACACAGCAAGCATAGTTGTG 115711

QY 2071 GATATGCGTAATTCGAAGTGAAGCTTCATCTGTATCATCGTCGCGGCGATGTGACATT 2130

Db 115712 GATATGCGTAATTCGAAGTGAAGCTTCATCTGTATCATCGTCGCGGCGATGTGACATT 115771

QY 2131 GAACCCGTACTTTAGAGTTGGAGATTACATCTCTACTCCAGAAATGTGCGTGAGCCG 2190
|||||
Db 115772 GAACCCGTACTTTAGAGTTGGAGATTACATCTCTACTCCAGAAATGTGCGTGAGCCG 115831

QY 2191 AAGAGTATGAGTATTATATGCGCTTTAAATATACGGCGGCTCTACAGCAAGTGCATC 2250
|||||
Db 115832 AAGAGTATGAGTATTATATGCGCTTTAAATATACGGCGGCTCTACAGCAAGTGCATC 115891

QY 2251 TCCATGTCCCGGTACTACAAAGCGTCCCGTCTCTGATTGAGTTGACCTAGCAAGCCG 2310
|||||
Db 115892 TCCATGTCCCGGTACTACAAAGCGTCCCGTCTCTGATTGAGTTGACCTAGCAAGCCG 115951

QY 2311 TTCTCTCTACTTCCCGAGAGTGCCTTTTACAGAGATCTCCAGCAATGACATTAATTC 2370
|||||
Db 115952 TTCTCTCTACTTCCCGAGAGTGCCTTTTACAGAGATCTCCAGCAATGACATTAATTC 116011

QY 2371 AAACCTACTCTTCTTACACTTCCCTCCAGACTCCGAGATTCTGTGCGCCCTCTCG 2430
|||||
Db 116012 AAACCTACTCTTCTTACACTTCCCTCCAGACTCCGAGATTCTGTGCGCCCTCTCG 116071

QY 2431 CATGCAACGCGGAGTGTGTTGAGAGCTGAACAAAGCAAGCCACAGCTGATGCGGCG 2490
|||||
Db 116072 CATGCAACGCGGAGTGTGTTGAGAGCTGAACAAAGCAAGCCACAGCTGATGCGGCG 116131

QY 2491 ACAGCACTGGCATTTACAGACATTTGAAACCTTCCGAGTCAAGAAAGTATATCT 2550
|||||
Db 116132 ACAGCACTGGCATTTACAGACATTTGAAACCTTCCGAGTCAAGAAAGTATATCT 116191

QY 2551 GGTCCCAAGACTCTTGTGTTAAATGCGAGGGGTGAATGCGCAAAAGTCCGCTCTTG 2610
|||||
Db 116192 GGTCCCAAGACTCTTGTGTTAAATGCGAGGGGTGAATGCGCAAAAGTCCGCTCTTG 116251

QY 2611 ATGCACACGTTAAGAATCGCAGATTAGCAGCCCTGTACAGACGAGCTCAGAGCT 2670
|||||
Db 116252 ATGCACACGTTAAGAATCGCAGATTAGCAGCCCTGTACAGACGAGCTCAGAGCT 116311

QY 2671 ATTCTGGGAATGCTGCAATGCGCAACAGCTTATGATTTCAATGACACCTTTTGGCA 2730
|||||
Db 116312 ATTCTGGGAATGCTGCAATGCGCAACAGCTTATGATTTCAATGACACCTTTTGGCA 116371

QY 2731 GAAGTGTATCAAAAAGAAAGGAAAGAAAGTGAACAGTATGATGCTTTTCTATCCAT 2790
|||||
Db 116372 GAAGTGTATCAAAAAGGAAAGGAAAGTGAACAGTATGATGCTTTTCTATCCAT 116431

QY 2791 GCCTGTACTTTTCAGCGGCTCCTTGGCAGACATCATAGTCAATTAATTAATGCTTGG 2850
|||||
Db 116432 GCCTGTACTTTTCAGCGGCTCCTTGGCAGACATCATAGTCAATTAATTAATGCTTGG 116491

QY 2851 CTAATTCATTTCTTTCCAAATGCTCTTAATGATTTGTACGTGACGCAAGCCAGATTC 2910
|||||
Db 116492 CTAATTCATTTCTTTCCAAATGCTCTTAATGATTTGTACGTGACGCAAGCCAGATTC 116551

QY 2911 TCCGTGAACCTGCAATGAGTGGCAATCACTGCACTGCGCTGCTCTCTCTCTCCGCC 2970
|||||
Db 116552 TCCGTGAACCTGCAATGAGTGGCAATCACTGCACTGCGCTGCTCTCTCTCTCCGCC 116611

QY 2971 TGCACCGTATGCGCGGCTTACAGATTTCTTTTAAATGAGTTGTACAGATCAGGT 3030
|||||
Db 116612 TGCACCGTATGCGCGGCTTACAGATTTCTTTTAAATGAGTTGTACAGATCAGGT 116671

QY 3031 AAAGTTCCTACAAGTATACAGAAAGTAGAACTTACCTGATCTTAACAGATCTCAT 3090
|||||
Db 116672 AAAGTTCCTACAAGTATACAGAAAGTAGAACTTACCTGATCTTAACAGATCTCAT 116731

QY 3091 TGAAGAGGAATATGCTTAACCTGTGCAATGAGCGTGAAGGAGGAGGAAAGACAGGACAA 3150
|||||
Db 116732 TGAAGAGGAATATGCTTAACCTGTGCAATGAGCGTGAAGGAGGAGGAAAGACAGGACAA 116791

QY 3151 GAAAGCTACATTTTAAAGTCTGTTATCTAGTGCAACATTAATTAACAGTCTTAAT 3210
|||||

Db 116792 GAAAGCTACATTTTAAAGCTCTGTTATCTAGTCACATTAATTAACAGTCTAATT 116851
QY 3211 GCACCTTATACCATGCTCTGCTGCTCCCAATCTGGCTCTTGGCTGTGTGCTGCTGGA 3270
Db 116852 GCACCTTATACCATGCTCTGCTGCTCCCAATCTGGCTCTTGGCTGTGTGCTGCTGGA 116911
QY 3271 CGCTTGAACGATCTTTGTGTGAGAAATCATGTTCTGACCTTGTGTCTACAAAGAGCCT 3330
Db 116912 CGCTTGAACGATCTTTGTGTGAGAAATCATGTTCTGACCTTGTGTCTACAAAGAGCCT 116971
QY 3331 TCTGGAACTGAGAGAAACATCTCTTGGCAATCTCTGACCACTCTCTCTACACATTT 3390
Db 116972 TCTGGAACTGAGAGAAACATCTCTTGGCAATCTCTGACCACTCTCTCTACACATTT 117031
QY 3391 TTTCTGACCTCCATCTCTTGGCTGCTGCTCTAGGAAATTTCTAGAGCTTCTTACT 3450
Db 117032 TTTCTGACCTCCATCTCTTGGCTGCTGCTCTAGGAAATTTCTAGAGCTTCTTACT 117091
QY 3451 ACTAATTCAGACAGCTCTCTCCCAAAAGCTGTTGACTAGTCTTCTAATGACCTTACATA 3510
Db 117092 ACTAATTCAGACAGCTCTCTCCCAAAAGCTGTTGACTAGTCTTCTAATGACCTTACATA 117151
QY 3511 TGTAGCATATATCTAATTAATTCATGTTCCAAATTAATTAATTAATTAATTAATTAAT 3570
Db 117152 TGTAGCATATATCTAATTAATTCATGTTCCAAATTAATTAATTAATTAATTAATTAAT 117211
QY 3571 CTGTTTCAAAAGTAAATGATGAGAGAGCTTTGATGATCTCTAATTTGTCACATTTCA 3630
Db 117212 CTGTTTCAAAAGTAAATGATGAGAGAGCTTTGATGATCTCTAATTTGTCACATTTCA 117271
QY 3631 GTCTCTCAATATCAGAGATCCCTAAGTCAGAGCTGCTAGTACAGAGCTTTTTCAGACTT 3690
Db 117272 GTCTCTCAATATCAGAGATCCCTAAGTCAGAGCTGCTAGTACAGAGCTTTTTCAGACTT 117331
QY 3691 CCTGCTTCTCAGCTCTTATATCTAGACACCGCATCATATCTCTAGAAATACAAAC 3750
Db 117332 CCTGCTTCTCAGCTCTTATATCTAGACACCGCATCATATCTCTAGAAATACAAAC 117391
QY 3751 TAATTCGACAGTGCAGATCCGACACCTGACCTGCTGAGGAGAGTGCAGACT 3810
Db 117392 TAATTCGACAGTGCAGATCCGACACCTGACCTGCTGAGGAGAGTGCAGACT 117451
QY 3811 TTGCTCTATTTACAAAAGAAAAGAAATTAACAACTTAAGCTCA 3857
Db 117452 TTGCTCTATTTACAAAAGAAAAGAAATTAACAACTTAAGCTCA 117498
RESULT 14
AF491814 29755 bp DNA linear PRI 20-MAR-2002
LOCUS AF491814 29755 bp DNA linear PRI 20-MAR-2002
DEFINITION Homo sapiens excision repair cross-complementing rodent repair
ACCESSION AF491814 deficiency, complementation group 4 (ERCC4) gene, complete cds.
VERSION AF491814.1 GI:19550956
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
REFERENCE 1 (bases 1 to 29755)
AUTHORS Rieder,M.J., Braun,A.C., Montoya,M.A., Chung,M.-W., Nguyen,C.P.,
Nguyen,D.A., Livingston,R.J., Poel,C.L., Robertson,P.D.,
Schackwitz,W.S., Sherwood,J.K., Wiltrik,L.A. and Nickerson,D.A.
COMMENT Submitted (12-MAR-2002) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
To cite this work please use: NIHES-SNPs, Environmental Genome
Project, NIHES ES15478, Department of Genome Sciences, Seattle, WA
(URL: <http://egp.gs.washington.edu>).
FEATURES
source
1..29755
/organism="Homo sapiens"
/db_xref="taxon:9606"

variation 122
/frequency="0.01"
/replace="c"
repeat_region 129..253
/rpt_family="L1"
/rpt_type-dispersed 248
/frequency="0.03"
/replace="g"
variation 424
/frequency="0.29"
/replace="t"
repeat_region 450..518
/rpt_family="Alu"
/rpt_type-dispersed 460
/frequency="0.01"
/replace="c"
variation 892
/frequency="0.01"
/replace="g"
variation 975
/frequency="0.01"
/replace="c"
variation 1038
/frequency="0.22"
/replace="a"
/gene="ERCC4"
join(<1068..1274,2934..3114,7469..7664,8937..9144,
11619..11799,13066..13194,15100..15210,16054..16651,
18674..18766,25632..25744,28524..29257)
/gene="ERCC4"
/product="excision repair cross-complementing rodent
repair deficiency, complementation group 4"
/protein_id="AL91593.1"
/db_xref="GI:19550957"
/translation="MESGOPARITAMAPLLEYEROLVLELDTDGLVYCARGLADRL
LVHEIOLHCPACLVIVINOPAREEYFNOLKIBGVENHPRVTNETISRSXYVT
QGSVTRTSRIILVVDPLTRIPSDLTITGLIVRAHRIISCOEAFILRLPROKNKGF
IKAFITNAVAEDTGFCHVERVMNLVFRKLYLWPRHVAANSFLEOHKRPVVEIIVSM
TPYMLAIOYALIDILNACLEKLNHNSLEVEDLSLENAIGKFPDRTIRYLDPLHQ
LGAKTKSLVODLKLRLTLQYLQYDCVFLNLESLRATKEAFQNSGWLFLDSSTS
MFANARVYHLDPDAKSKREKISEKMEIKEGEETKEKLELVESNPKMEALVEYKETE
AENKESDALGGPGCVILICASDORTCSOLRYITLGADEFLRLRYRTPEKDSAAEEWM
MKRKSDSSKRIRKSHKRPDPONKERASIKERTLKKRKLLTOMVCKPELEEDG
DVEGTRREIISSESCPEIKHEEDVNLSSDAHGLAEPLTIHPLHPCSDPAL
TRVLHEVPRYVLYDAELTFVRQLEIYRASRGKPLRYFLYGGSTEOURLTLAR
KEKEAFELKIREKASWVPERGERDETINLDVRCYASADVSTTRKAGQEQNGQO
SIVDMRSESESLPLIHRGIDIEPVTLGVGYLIIPKCVKRSISDILGSLNNGR
LYSQCISMSRYKRPVLLIPEFDSKPSLRSRALLOEISNSISKSLTLLTLPFRL
RLWCSPHATLAFELIKOSKROPAAATLATASETLPSEKYNPGQDDLLMP
GVNAKNCRSLMHVKNIAELASODELITSLGMAANAKOLYFIHTSPAEVYSKRG
KR"
1100
/gene="ERCC4"
/frequency="0.01"
/replace="t"
variation 1323
/gene="ERCC4"
/frequency="0.15"
/replace="a"
variation 1353
/gene="ERCC4"
/frequency="0.01"
/replace="c"

Db 29045 GGTCCCCAGACTCTTCTTTAAATGCGAGGGGTGATGCCAAAAGTCCCGCTCTTG 29104
Qy 2611 ATGCACCACTTATAGAACATCGCAGATTAGACGCCCTGTACAGAGAGACTCAGAGT 2670
Db 29105 ATGCACCACTTATAGAACATCGCAGATTAGACGCCCTGTACAGAGAGACTCAGAGT 29164
Qy 2671 ATTCTGGGAATGCTGCAATGCGAATGAGCTTTATGATTTCACTTACACCTTTTGA 2730
Db 29165 ATTTGGGAATGCTGCAATGCGAATGAGCTTTATGATTTCACTTACACCTTTTGA 29224
Qy 2731 GAAGTGTATCAAAAGGAAAGGAAAGTGAACAGTATGAGCTTTTCTTATCCAT 2790
Db 29225 GAAGTGTATCAAAAGGAAAGGAAAGTGAACAGTATGAGCTTTTCTTATCCAT 29284
Qy 2791 GCTGTACTTTTTCAGCGGCTCCTTGGCAGACATCATAGTCTATTATTTATTTG 2850
Db 29285 GCTGTACTTTTTCAGCGGCTCCTTGGCAGACATCATAGTCTATTATTTATTTG 29344
Qy 2851 CTATTCTATTCTTTTCAATGCTCTTAATGATTTGTAAGGAGACAGAAAGCAGATTC 2910
Db 29345 CTATTCTATTCTTTTCAATGCTCTTAATGATTTGTAAGGAGACAGAAAGCAGATTC 29404
Qy 2911 TCTCTGACTGTGCACTTGAAGCATCTGACCTGTGCTGCTCTTTTCTCTCC 2970
Db 29405 TCTCTGACTGTGCACTTGAAGCATCTGACCTGTGCTGCTCTTTTCTCTCC 29464
Qy 2971 TGACCGCTCTATGCGGGCTTACATGTTTCTTTTAAATGAGTTTGCAGATCAGT 3030
Db 29465 TGACCGCTCTATGCGGGCTTACATGTTTCTTTTAAATGAGTTTGCAGATCAGT 29524
Qy 3031 AAGTTCTCTACAGATTTACAGAGATGAGAACTTACCTGATCTTACAGATCTCAT 3090
Db 29525 AAGTTCTCTACAGATTTACAGAGATGAGAACTTACCTGATCTTACAGATCTCAT 29584
Qy 3091 TAGAAGGAATATGCTAAGCTTGCATGAGCGGTGCAAGGAGGAAAGAGCGACAA 3150
Db 29585 TAGAAGGAATATGCTAAGCTTGCATGAGCGGTGCAAGGAGGAAAGAGCGACAA 29644
Qy 3151 GAAGGTACCATTTTAAAGCTCTTGTATCTAGTGCAATTAATTAACAGCTTAAT 3210
Db 29645 GAAGGTACCATTTTAAAGCTCTTGTATCTAGTGCAATTAATTAACAGCTTAAT 29704
Qy 3211 GCACTTATACCATGTCTGTGCTCTCAAAATCTGCTCTTGTCTGTG 3261
Db 29705 GCACTTATACCATGTCTGTGCTCTCAAAATCTGCTCTTGTCTGTG 29755

RESULT 15
ARI83108 1101 bp DNA linear PAT 20-APR-2002
LOCUS ARI83108
DEFINITION Sequence 12 from patent US 6340566.
ACCESSION ARI83108
VERSION ARI83108.1 GI:20226701
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1101)
AUTHORS McCutchen-Maloney,S.L.
TITLE Detection and quantification of single nucleotide polymorphisms. DNA
JOURNAL sequence variations, DNA mutations, DNA damage and DNA mismatches
FEATURES
source Location/Qualifiers
1..1101
BASE COUNT 324 a 230 c 246 g 301 t
ORIGIN
Query Match 28.5%; Score 1101; DB 6; Length 1101;
Best Local Similarity 100.0%; Pred. No. 4.9e-251;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 ATGGCGCGCTGCTGAGTACGAGCAGACACTGTGCTGGAACGCTGACACTGACGGG 105

Db 1 ATGGCGCGCGCTGCTGAGTACGAGCAGACACTGTGCTGGAACGCTGACACTGACGGG 60
Qy 106 CTAGTAGTGTGGCCCGCGGGCTGCGCGGAGCGAGCTCTCTTACCACTTTCTCCAGCTG 165
Db 61 CTAGTAGTGTGGCCCGCGGGCTGCGCGGAGCGAGCTCTCTTACCACTTTCTCCAGCTG 120
Qy 166 CACTGCGACCCAGCGCTGGTGTGCTGCTCAACAGCAGCGCGCGAGAGAGAT 225
Db 121 CACTGCGACCCAGCGCTGGTGTGCTGCTCAACAGCAGCGCGCGAGAGAGAT 180
Qy 226 TTTATCAATCAGCTGAAGATAGAAAGAGTTGAACACCTCCCTGCGGTGTAAACAAATGA 285
Db 181 TTTATCAATCAGCTGAAGATAGAAAGAGTTGAACACCTCCCTGCGGTGTAAACAAATGA 240
Qy 286 ATCAAGCAACAGCTGCTATGAAAGTTTACACCAAGGAGTGTATTTATTTGGACAAGT 345
Db 241 ATCAAGCAACAGCTGCTATGAAAGTTTACACCAAGGAGTGTATTTATTTGGACAAGT 300
Qy 346 AGGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 405
Db 301 AGGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 406 TTGGTGTATAGACCCCAAGCAATATGAGTCTTGTCAAGACATTCATCTTGGCCTC 465
Db 361 TTGGTGTATAGACCCCAAGCAATATGAGTCTTGTCAAGACATTCATCTTGGCCTC 420
Qy 466 TTTCGCGCAAGAAACAACCTGTTTATTAATGAAGCTTTCACACANVGTGTGCTT 525
Db 421 TTTCGCGCAAGAAACAACCTGTTTATTAATGAAGCTTTCACACANVGTGTGCTT 480
Qy 526 GATACGTGTTTGTGATGTGAAGAGTGAAGTGAAGTCTTTTGTGAGAACTGTAT 585
Db 481 GATACGTGTTTGTGATGTGAAGAGTGAAGTGAAGTCTTTTGTGAGAACTGTAT 540
Qy 586 CTGTGGCCAAAGTTCCATGATGACGATTAATCTATTTTGAACAGCAACAACTGAAGT 645
Db 541 CTGTGGCCAAAGTTCCATGATGACGATTAATCTATTTTGAACAGCAACAACTGAAGT 600
Qy 646 GTAGAAATCCATGTTTCTATGACACCTGACCTGCTTATGACAGCTGTATACGTGAG 705
Db 601 GTAGAAATCCATGTTTCTATGACACCTGACCTGCTTATGACAGCTGTATACGTGAG 660
Qy 706 ATTTTAAATCATGTCTTAAAGAACTTAAATGCCATTAACCATCGCTTGAAGTGAAGAT 765
Db 661 ATTTTAAATCATGTCTTAAAGAACTTAAATGCCATTAACCATCGCTTGAAGTGAAGAT 720
Qy 766 TTTATCTTTAGAAAATGCTATGAGAAACCTTTGACAAAGCATCCGCCATTTCTGGAT 825
Db 721 TTTATCTTTAGAAAATGCTATGAGAAACCTTTGACAAAGCATCCGCCATTTCTGGAT 780
Qy 826 CTTTGTGGACACGCTGAGGCAACGATTAATCTTATGAGCAAGTTTGAAGATTTA 885
Db 781 CTTTGTGGACACGCTGAGGCAACGATTAATCTTATGAGCAAGTTTGAAGATTTA 840
Qy 886 CGAATCTGTGCAAGTATCTCTCACTATGATGTGTCACATTTCTTAACTTCTGGAA 945
Db 841 CGAATCTGTGCAAGTATCTCTCACTATGATGTGTCACATTTCTTAACTTCTGGAA 900
Qy 946 TCTGTGAGACCAAGGAAAGCTTTTGTGACAGATTCAGGTTGCTTTCTTGAATCC 1005
Db 901 TCTGTGAGACCAAGGAAAGCTTTTGTGACAGATTCAGGTTGCTTTCTTGAATCC 960
Qy 1006 AGCAGCTGATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCAAAATG 1065
Db 961 AGCAGCTGATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCAAAATG 1020
Qy 1066 AGTAAAAAGAAAAATATGTGAAAAATGAATTAAGAGAGGAGAAACAAAAAG 1125
Db 1021 AGTAAAAAGAAAAATATGTGAAAAATGAATTAAGAGAGGAGAAACAAAAAG 1080
Qy 1126 GAATGCTCTAGAAACAAC 1146

Wed Nov 27 08:49:07 2002

us-09-919-497-14.rge

Page 25

Db 1081 GAACTGGTCTAGAAAGCAAC 1101

Search completed: November 26, 2002, 22:46:38
Job time : 8482.76 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:42:45 : Search time 540.781 Seconds
(without alignments)
16061.850 Million cell updates/sec

Title: US-09-919-497-14

Perfect score: 3857
Sequence: 1 ggaagagcttcacatggagtc.....agaatacaacctaagctca 3857

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapect 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: N_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3857	100.0	3857	24	ABK35494
2	2875	74.5	2890	18	AA197897
3	2857.4	74.1	2900	21	AA16235
4	2718	70.5	2718	22	AA563230
5	1829.4	47.4	35641	24	ABN95727
6	1829.4	47.4	35641	24	ABN64428
7	1101	28.5	1101	22	AA563229
8	256.2	6.6	601	24	ABL87494
9	253.2	6.6	2995	23	ABL06559

10	159	4.1	5313	23	ABL06558	Drosophila melanog
C 11	154.6	4.0	3857	24	ABK35494	Human endometrial
C 12	154.6	4.0	35641	24	ABN95727	Gene #2225 used to
C 13	154.6	4.0	35641	24	ABN64428	Stomach cancer rel
C 14	135.6	3.5	574	24	ABQ21888	Oligonucleotide fo
C 15	135.6	3.5	574	24	ABQ21889	Oligonucleotide fo
C 16	135.6	3.5	574	24	ABQ32918	Oligonucleotide fo
C 17	135.6	3.5	574	24	ABQ32919	Oligonucleotide fo
C 18	135.6	3.5	6866	24	ABL49319	Human polynucleoti
C 19	135.6	3.5	6866	24	ABL52666	Human immune syste
C 20	131.8	3.4	574	24	ABQ21886	Oligonucleotide fo
C 21	131.8	3.4	574	24	ABQ21887	Oligonucleotide fo
C 22	131.8	3.4	574	24	ABQ32920	Oligonucleotide fo
C 23	131.8	3.4	574	24	ABQ32921	Oligonucleotide fo
C 24	131.8	3.4	6866	24	ABL49320	Human polynucleoti
C 25	131.8	3.4	6866	24	ABL52667	Human immune syste
C 26	112.4	2.9	589	21	AA197599	Fusarium venenatum
C 27	111.4	2.9	842	23	ABV03504	Human prostate exp
C 28	85.4	2.2	405	21	AAH30546	Human colon cancer
C 29	80	2.1	430	24	ABO55897	Human ovarian anti
C 30	79.4	2.1	570	23	ABV22282	Human prostate exp
C 31	79.4	2.1	570	23	ABV28116	Human prostate exp
C 32	79.4	2.1	605	23	ABV33803	Human prostate exp
C 33	79.4	2.1	605	23	ABV42696	Human prostate exp
C 34	79.4	2.1	871	23	ABV12673	Human prostate exp
C 35	72.2	1.9	288	20	AAV89510	EST clone COL151.
C 36	72	1.9	110	21	AA12928	Human secreted pro
C 37	70.6	1.8	318	22	ABAI6642	Human nervous syst
C 38	70.6	1.8	6318	22	AA199118	Human excretory re
C 39	70.6	1.8	6318	22	AA163468	Human kidney relat
C 40	70.6	1.8	32127	22	AA199255	Human excretory re
C 41	70.6	1.8	32127	22	AA163605	Human kidney relat
C 42	70.6	1.8	72215	22	AAK68832	Human immune/thema
C 43	70.6	1.8	236303	22	AA511614	Human genomic DNA
C 44	70.2	1.8	9566	22	ABR20411	Human nervous syst
C 45	70.2	1.8	9566	22	AA137121	Human musculoskele

ALIGNMENTS

```

RESULT 1
ABK35494
ID ABK35494 standard; DNA: 3857 BP.
XX
AC ABK35494:
XX
DT 08-MAY-2002 (first entry)
XX
DE Human endometrial cancer related gene, ERCC4.
XX
KW Human; ds: gene; endometrial cancer; differential expression;
KW DNA microarray; protein microarray.
XX
OS Homo sapiens.
XX
PN MO200209573-A2.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US24104.
XX
PR 31-JUL-2000; 2000US-221735P.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Mutter GL:
XX
DR WPI: 2002-179967/23.
XX
PT P-PSDB: AAU84274.
XX
PT Diagnosing endometrial cancer comprises determining expression of
nucleic acid molecules or expression products that are differentially

```

PT expressed in normal and malignant endometrium -
XX
PS Claim 1: Page 64-66; 233pp; English.
XX
CC The invention relates to diagnosing endometrial cancer in a subject
CC suspected of having endometrial cancer comprising determining the
CC expression of a set of nucleic acid molecules or expression products in
CC an endometrial sample suspected of being cancerous, where the set of
CC nucleic acid molecules comprises at least 2 nucleic acid molecules
CC selected from 50 fully defined sequences as given in the specification.
CC The nucleic acids are used as an array of at least 2 of the 50
CC nucleic acids bound to a solid substrate. Also included is a solid-phase
CC protein microarray comprising at least 2 antibodies or its antigen
CC binding fragments, that specifically bind at least 2 different
CC polypeptides from the 50 fully defined sequences as given in the
CC specification, fixed to a solid substrate. The methods and arrays are
CC useful for the diagnosis of endometrial cancer, selecting and monitoring
CC treatment regimes and identification of lead compounds useful for the
CC treatment of endometrial cancer. The present sequence is one of 50
CC genes differentially expressed between cancerous and non-cancerous
CC samples.
XX
SQ Sequence 3857 BP; 1134 A; 851 C; 850 G; 1022 T; 0 other.

Query Match 100.0%; Score 3857; DB 24; Length 3857;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGACCTTCATGAGTCAGGCGCGGCTCGACGCGATTGCCATGGCCGCTGCTG 60
DB 1 GGAAGACCTTCATGAGTCAGGCGCGGCTCGACGCGATTGCCATGGCCGCTGCTG 60
QY 61 GACTACAGAGCAGACGCTGGTGTCTGAACCTCTGACACCTGACGCGGCTAGTGTGCGCC 120
DB 61 GACTACAGAGCAGACGCTGGTGTCTGAACCTCTGACACCTGACGCGGCTAGTGTGCGCC 120
QY 121 CGGCGGCTCGGCGCGGACCGGCTCTCTACCACTTCTCAGCTGACCTGCCACCGACGC 180
DB 121 CGGCGGCTCGGCGCGGACCGGCTCTCTACCACTTCTCAGCTGACCTGCCACCGACGC 180
QY 181 TGGCTGTGCTGTGCTGCTCAACACGCGCGGAGAGAGTATTTTATCATCAGCTG 240
DB 181 TGGCTGTGCTGTGCTGCTCAACACGCGCGGAGAGTATTTTATCATCAGCTG 240
QY 241 AAGATGAAGAGAGTGAACACCTCCCTCGCGCTGAACAAATGAATACCAACAGT 300
DB 241 AAGATGAAGAGAGTGAACACCTCCCTCGCGCTGAACAAATGAATACCAACAGT 300
QY 301 CGCTATGAAGTTTACACACAGGTGTTATATTTGGCAGAGTAGATCTTGTGTT 360
DB 301 CGCTATGAAGTTTACACACAGGTGTTATATTTGGCAGAGTAGATCTTGTGTT 360
QY 361 GACTTGTACGTAGTAACTTACCTGAGATTTATTTACGTGCTTGTGTTATACGCC 420
DB 361 GACTTGTACGTAGTAACTTACCTGAGATTTATTTACGTGCTTGTGTTATACGCC 420
QY 421 CACAGATATATCGAGTCTTGTCAAGAAGCATTCATCTTGCGCCTCTTGGCCGAAAAAC 480
DB 421 CACAGATATATCGAGTCTTGTCAAGAAGCATTCATCTTGCGCCTCTTGGCCGAAAAAC 480
QY 481 AAACGTGTTTTTATTAAGCTTTCACAGACATGCTGTTGCCCTTGTATGACTGTTTTGT 540
DB 481 AAACGTGTTTTTATTAAGCTTTCACAGACATGCTGTTGCCCTTGTATGACTGTTTTGT 540
QY 541 CATGTGAAAGAGTGAATCTTTTGTGAGAGAACTGTATCTGTGCGCAAGGTTTC 600
DB 541 CATGTGAAAGAGTGAATCTTTTGTGAGAGAACTGTATCTGTGCGCAAGGTTTC 600
QY 601 CATGTGACAGTAACTCATTTTGTAGAACAGACAACTGAAGTGTAGAAATCATGTT 660
DB 601 CATGTGACAGTAACTCATTTTGTAGAACAGACAACTGAAGTGTAGAAATCATGTT 660
QY 661 TCTATGACACCTACATGCTTGTCTATACAGACTGTATACGTATTTAAATGATGT 720

DB 661 TCTATGACACCTACATGCTTGTCTATACAGACTGTATACGTATTTAAATGATGT 720
QY 721 CTAAGGAACTAAATATGCCATTAACCATCGCTGAAGTGGAAATTTATCTTTAGAAAT 780
DB 721 CTAAGGAACTAAATATGCCATTAACCATCGCTGAAGTGGAAATTTATCTTTAGAAAT 780
QY 781 GCATTGGAAGACCTTTTGCAGACAAATCCGCATTATCTGATCTTTGTGGACAC 840
DB 781 GCATTGGAAGACCTTTTGCAGACAAATCCGCATTATCTGATCTTTGTGGACAC 840
QY 841 CTTGAGCCAAAGACTAATCTTAGTTCAAGATTTGAAGATATTACGACTTGTGCAG 900
DB 841 CTTGAGCCAAAGACTAATCTTAGTTCAAGATTTGAAGATATTACGACTTGTGCAG 900
QY 901 TATCTCTCTGATGATGATGTTGTGCATATTTCTTAATCTTGTGAATCTCTGAGACAC 960
DB 901 TATCTCTCTGATGATGATGTTGTGCATATTTCTTAATCTTGTGAATCTCTGAGACAC 960
QY 961 GAAAAAGCTTTTGTGCAAGATTCAGGTGGCTGTTCTTGACCTCCAGCACCTGATGTTT 1020
DB 961 GAAAAAGCTTTTGTGCAAGATTCAGGTGGCTGTTCTTGACCTCCAGCACCTGATGTTT 1020
QY 1021 ATAAATGCTCGACCAAGGTTTATCATCTTCCAGATGCCAAATGAGTAAAGAAAAA 1080
DB 1021 ATAAATGCTCGACCAAGGTTTATCATCTTCCAGATGCCAAATGAGTAAAGAAAAA 1080
QY 1081 ATATCTGAAAAAATGGAATTTAAGAGGGAAGAAACAAAGAACTGCTCTAGAA 1140
DB 1081 ATATCTGAAAAAATGGAATTTAAGAGGGAAGAAACAAAGAACTGCTCTAGAA 1140
QY 1141 AGCAACCCAAAGTGGAGGACACTGACTGAATTTTAAAGAAATTTAGGACAGAAAAATAG 1200
DB 1141 AGCAACCCAAAGTGGAGGACACTGACTGAATTTTAAAGAAATTTAGGACAGAAAAATAG 1200
QY 1201 GAGAGTGAAGCTTGTGTGCTCAGTCAATGATTTTGTGCAAGTGTGACCGAACA 1260
DB 1201 GAGAGTGAAGCTTGTGTGCTCAGTCAATGATTTTGTGCAAGTGTGACCGAACA 1260
QY 1261 TGTTCCTCAGCTGAGACATATATCATCTTGTGAGCGGAGGCTTCTTATTTGAGGCTCTAC 1320
DB 1261 TGTTCCTCAGCTGAGACATATATCATCTTGTGAGCGGAGGCTTCTTATTTGAGGCTCTAC 1320
QY 1321 AGGAAAACTTTGAGAAGATAGCAAGCTGAAGAAGTGTGATGAATTTAGAGAGAA 1380
DB 1321 AGGAAAACTTTGAGAAGATAGCAAGCTGAAGAAGTGTGATGAATTTAGAGAGAA 1380
QY 1381 GACAGTTCAAGAGATTTAGGAATCTCACAAAAGACTTAAGACCCCAAAACAAAGAA 1440
DB 1381 GACAGTTCAAGAGATTTAGGAATCTCACAAAAGACTTAAGACCCCAAAACAAAGAA 1440
QY 1441 CGGCGCTTTCACCAAGAAAGAACCTCAAAAAGAAAAACGGAAGTGAACCTTAACCTCAA 1500
DB 1441 CGGCGCTTTCACCAAGAAAGAACCTCAAAAAGAAAAACGGAAGTGAACCTTAACCTCAA 1500
QY 1501 ATGTTAGAAAAACCTGAGAACTGGAAGAGAGAGATGTGAGAGAGATATCTGTCA 1560
DB 1501 ATGTTAGAAAAACCTGAGAACTGGAAGAGAGAGATGTGAGAGAGATATCTGTCA 1560
QY 1561 GAAATTAAGCAGTAGCCAGAAAGCTCCCGGAAGAAATTTAGCATAGAAATTTGATGTA 1620
DB 1561 GAAATTAAGCAGTAGCCAGAAAGCTCCCGGAAGAAATTTAGCATAGAAATTTGATGTA 1620
QY 1621 AATTTGCAATCGGATGCTGTTTGGAAATCTGAAAGAACCCCTCAATCATCATCCG 1680
DB 1621 AATTTGCAATCGGATGCTGTTTGGAAATCTGAAAGAACCCCTCAATCATCATCCG 1680
QY 1681 CTTCTGGTTGACAGCAACCCCTATGCTGTGACAAGGCTACTACATGAGTGAAGCCACA 1740
DB 1681 CTTCTGGTTGACAGCAACCCCTATGCTGTGACAAGGCTACTACATGAGTGAAGCCACA 1740
QY 1741 TACGTGTTCTTTATGACGACAGCTTAACCTTTGTTGGCAGCTTCAAAATTTACAGCGG 1800

D 1741 TACGTGGTCTTTATGACGACGACCTAACCTTTGTGGCAGCTGTAATTTACAGGGCG 1800
Q 1801 AGTAGGCGTGGGAACCTCTGAGGGTTACTTCTTATATACGAGGTTCACGAGAA 1860
D 1801 AGTAGGCGTGGGAACCTCTGAGGGTTACTTCTTATATACGAGGTTCACGAGAA 1860
Q 1861 CAAAGCATATCATGCTGCTTGGGAAAGAAAGAAAGCTTTTGAAGAACTATAGGAA 1920
D 1861 CAAAGCATATCATGCTGCTTGGGAAAGAAAGAAAGCTTTTGAAGAACTATAGGAA 1920
Q 1921 AAAGCAAGCATGTTGTCCTGGAAGAAAGAGAGGAGAGATGAACAACTTAGACCTA 1980
D 1921 AAAGCAAGCATGTTGTCCTGGAAGAAAGAGAGGAGAGATGAACAACTTAGACCTA 1980
Q 1981 GTAGAGGACAGCATGCTGAGATGTTTCCACTGACACTGGAAGAGCGGTGGCCAGAA 2040
D 1981 GTAGAGGACAGCATGCTGAGATGTTTCCACTGACACTGGAAGAGCGGTGGCCAGAA 2040
Q 2041 CAGAATGGTACACAGCAAAAGCATAGTTGTGATATGCGTGAATTTGCAAGTGAAGTTC 2100
D 2041 CAGAATGGTACACAGCAAAAGCATAGTTGTGATATGCGTGAATTTGCAAGTGAAGTTC 2100
Q 2101 TCTCTGATCCATCGTGGGGGATTTGACATTTAACCCTGACTTTAGAGCTTGAGATTAC 2160
D 2101 TCTCTGATCCATCGTGGGGGATTTGACATTTAACCCTGACTTTAGAGCTTGAGATTAC 2160
Q 2161 ATCTCATCTCAGAAATGTGGTGAGGCGCAAGAGTATCAGTGAATTTAATGGGCTTTA 2220
D 2161 ATCTCATCTCAGAAATGTGGTGAGGCGCAAGAGTATCAGTGAATTTAATGGGCTTTA 2220
Q 2221 AATAAGGCGCCCTCTACAGCCAGTGCATCTCATGTCCTCCCTACTACAGCGTCCGCTG 2280
D 2221 AATAAGGCGCCCTCTACAGCCAGTGCATCTCATGTCCTCCCTACTACAGCGTCCGCTG 2280
Q 2281 CTCTGATGATGTTGACCCCTAGCAAGCCCTTCTCTCACTTCCCGAGGCTGCTTGT 2340
D 2281 CTCTGATGATGTTGACCCCTAGCAAGCCCTTCTCTCACTTCCCGAGGCTGCTTGT 2340
Q 2341 CAGAGATCTCTCAGCAATGATATTAGTTCAGAACTCACTCTTCTACACTTCCGCTGCT 2400
D 2341 CAGAGATCTCTCAGCAATGATATTAGTTCAGAACTCACTCTTCTACACTTCCGCTGCT 2400
Q 2401 AGACTAGGATCTCTGTGTCGCCCTCTCTCATGCAAGCGGCGAGTTTGGAGAGCTG 2460
D 2401 AGACTAGGATCTCTGTGTCGCCCTCTCTCATGCAAGCGGCGAGTTTGGAGAGCTG 2460
Q 2461 AAACAAAGCAAGCCAGCCGATGATGCGCGACAGCATGCGCATTTACAGAGATTTGAA 2520
D 2461 AAACAAAGCAAGCCAGCCGATGATGCGCGACAGCATGCGCATTTACAGAGATTTGAA 2520
Q 2521 ACCCTCCGAGTCAAGAGATTAATGCTGTGCCCAAGCTCTGTAAATAATGGCA 2580
D 2521 ACCCTCCGAGTCAAGAGATTAATGCTGTGCCCAAGCTCTGTAAATAATGGCA 2580
Q 2581 GGGGTGATGCAAAAACCTGCGCTCTTGTGATGACACACGTTAAGAACATGCGAGAA 2640
D 2581 GGGGTGATGCAAAAACCTGCGCTCTTGTGATGACACACGTTAAGAACATGCGAGAA 2640
Q 2641 GCAGCCCTGTCAAGAGAGAGCTCAAGATATTTCTGGGGAATGCTGCAAAATGCCAAAC 2700
D 2641 GCAGCCCTGTCAAGAGAGAGCTCAAGATATTTCTGGGGAATGCTGCAAAATGCCAAAC 2700
Q 2701 CTTATATGATTTCAATTCACACTCTTTTGCAGAAAGTCAAAAAGGGAAGGAAAAAG 2760
D 2701 CTTATATGATTTCAATTCACACTCTTTTGCAGAAAGTCAAAAAGGGAAGGAAAAAG 2760
Q 2761 TGAACAGATGAGGCTGTTTCTTATCCCATGCTGACTTTTACGCGGCTCCTTGGCAGA 2820
D 2761 TGAACAGATGAGGCTGTTTCTTATCCCATGCTGACTTTTACGCGGCTCCTTGGCAGA 2820
Q 2821 CATCATAGATGATTAATTAATGTTGTTGCTAATTTCACTTCTTCCAAATGCTTAA 2880
D 2821 CATCATAGATGATTAATTAATGTTGTTGCTAATTTCACTTCTTCCAAATGCTTAA 2880

Q 2881 ATTGTAGCGTGACAGAGAGCCAGGATTCCTCTGAACTCTGCAAGTATGAGCATCACTTG 2940
D 2881 ATTGTAGCGTGACAGAGAGCCAGGATTCCTCTGAACTCTGCAAGTATGAGCATCACTTG 2940
Q 2941 AACTGTGCTGTGCGCTCTCTTTTCTCCCTGACACCGCTCTATGCGGCTTACGATGTT 3000
D 2941 AACTGTGCTGTGCGCTCTCTTTTCTCCCTGACACCGCTCTATGCGGCTTACGATGTT 3000
Q 3001 CTTTAAATAGAGTTTGTGAGATCAGGTAAAGTTCTTACAAAGTATACAGAAAGTAG 3060
D 3001 CTTTAAATAGAGTTTGTGAGATCAGGTAAAGTTCTTACAAAGTATACAGAAAGTAG 3060
Q 3061 AAACCTTCTGATCTCAAGATCTCATTTAGAAAGGATATGCTAAGCCCTGGCATGA 3120
D 3061 AAACCTTCTGATCTCAAGATCTCATTTAGAAAGGATATGCTAAGCCCTGGCATGA 3120
Q 3121 CGGTGACAGGAGGGAAGAGACAGACAGCAAGAAACCTCAATTTTAAACAGTCTTGT 3180
D 3121 CGGTGACAGGAGGGAAGAGAGACAGACAGCAAGAAACCTCAATTTTAAACAGTCTTGT 3180
Q 3181 TCTAGTCAACATTAATTAACAGTCTTAATGCACTTATACCATTTCTGCTGGCTTCCA 3240
D 3181 TCTAGTCAACATTAATTAACAGTCTTAATGCACTTATACCATTTCTGCTGGCTTCCA 3240
Q 3241 AATCGGCTCTTGTGCTGCTGCTGCTGCTGAGCGCTTGAACATGATGTTGTAGGAATCA 3300
D 3241 AATCGGCTCTTGTGCTGCTGCTGCTGCTGAGCGCTTGAACATGATGTTGTAGGAATCA 3300
Q 3301 TGTCTGACCCCTTGTCTCAAAAGAGCCCTCTGGAACCTGAGAAAGACATCTTGTG 3360
D 3301 TGTCTGACCCCTTGTCTCAAAAGAGCCCTCTGGAACCTGAGAAAGACATCTTGTG 3360
Q 3361 CCATTCCTGACAGTGTCTCTACACATTTTCTTCACTGCTCATACCTCTGCTGCTG 3420
D 3361 CCATTCCTGACAGTGTCTCTACACATTTTCTTCACTGCTCATACCTCTGCTGCTG 3420
Q 3421 TCTAAGAAATTTCAATGAGGCTTCTCTACTAATTTGAAGACAGTCTCTCAAAAAGT 3480
D 3421 TCTAAGAAATTTCAATGAGGCTTCTCTACTAATTTGAAGACAGTCTCTCTCAAAAAGT 3480
Q 3481 GTTGAATGCTCTTAATGACCTTAACATATGATAGCATATATTAATTTCAATTTGCA 3540
D 3481 GTTGAATGCTCTTAATGACCTTAACATATGATAGCATATATTAATTTCAATTTGCA 3540
Q 3541 AATTAGTATTTTAAAGCAAAATGATACGTTTGGAAAAAGTTAATGATGAAGACT 3600
D 3541 AATTAGTATTTTAAAGCAAAATGATACGTTTGGAAAAAGTTAATGATGAAGACT 3600
Q 3601 CTAGAAATTCGAATTTTGGACATATTCAGTCTCTAATATCAGATATCCCTAAGTCA 3660
D 3601 CTAGAAATTCGAATTTTGGACATATTCAGTCTCTAATATCAGATATCCCTAAGTCA 3660
Q 3661 GCTGGCTAGTTACAGAGTTTTTTCAGACTCTCTGTTTCTGAGCTTATATCTTAAGAC 3720
D 3661 GCTGGCTAGTTACAGAGTTTTTTCAGACTCTCTGTTTCTGAGCTTATATCTTAAGAC 3720
Q 3721 ACCAGCATATATCTCTAGAAATTAACAACCTAATTTGCAAGTGAAGCGGAGATGCACT 3780
D 3721 ACCAGCATATATCTCTAGAAATTAACAACCTAATTTGCAAGTGAAGCGGAGATGCACT 3780
Q 3781 GCACCCCTGCTGGGCGACAGAGTGAAGCTTGTCTCTATTAACAAGAAAAAGAAAGA 3840
D 3781 GCACCCCTGCTGGGCGACAGAGTGAAGCTTGTCTCTATTAACAAGAAAAAGAAAGA 3840
Q 3841 AATACACCTAAGCTCA 3857
D 3841 AATACACCTAAGCTCA 3857

RESULT 2
AAT97897
ID AAT97897 standard; cdna: 2890 bp.

XX	AA197897;
AC	
DT	11-MAY-1998 (first entry)
XX	
DE	Human DNA repair enzyme RAD cDNA.
XX	
KM	RAD; DNA repair enzyme; human; xeroderma pigmentosum; colon cancer;
KW	hereditary nonpolyposis coli; HNPCC; diagnosis; gene therapy; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	25..2742
FT	/+tag= a
PN	MOJ9742209-A1.
PD	13-NOV-1997.
PF	03-MAY-1996; 96WO-US06221.
PR	03-MAY-1996; 96WO-US06221.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(ICRF-) ICRF HALL LAB CLARE.
PA	(UYRO-) UNIV ROTTERDAM ERASMUS.
PI	Arliza RR, Hoelmakers JHU, Ruben SM, Sijbers A;
PI	Weil YF, Wood RD;
DR	WPI: 1997-558900/51.
DR	P-PsDB: AAM36879.
XX	
PT	Polynucleotide encoding human DNA repair enzyme, RAD - useful to
PT	develop products to diagnose and treat, e.g. xeroderma pigmentosum
PT	and colon cancer
PS	Claim 4: Page 60-61; 81pp; English.
XX	
CC	This cDNA sequence includes a coding region for the human DNA
CC	repair enzyme RAD (see AAM36879). Degenerate primers based on the
CC	homology between Saccharomyces cerevisiae RAD1, Schizosaccharomyces
CC	pombe Rad16+ and Drosophila melanogaster mei-9 gene products were
CC	used in a RT-PCR with total HeLa cell RNA to amplify a fragment
CC	encoding the homologous human gene. The PCR product was used as a
CC	probe to identify a positive clone from a human testis cDNA
CC	library. The sequences were then used to obtain the full-length
CC	human RAD cDNA. The isolated polynucleotide can be used for the
CC	recombinant production of RAD polypeptides in host cells. Methods
CC	are disclosed for assessing RAD expression in cells by determining
CC	RAD polypeptides or RAD-encoding mRNA, for treating diseases
CC	associated with defective DNA repair such as xeroderma pigmentosum,
CC	colon cancer and hereditary nonpolyposis coli (HNPCC) in vitro, ex
CC	vivo or in vivo by exposing cells to RAD polypeptides or
CC	polynucleotides, for assaying genetic variation and aberrations in
CC	RAD genes, and for administering a RAD polypeptide or polynucleotide
CC	to an organism to augment RAD function or remediate RAD dysfunction.
XX	
SQ	Sequence 2890 BP; 862 A; 630 C; 672 G; 726 T; 0 other:
	Query Match 74.5%; Score 2875; DB 18; Length 2890;
	Best Local Similarity 99.8%; Pred No. 0;
	Matches 2878; Conservative 0; Mismatches 5; Indels 0; Gaps
OY	22 GGGCAGCGGCTGCAGCGATTGCCATGCGCGCTGTGACTACGACGACAGCTGTG 81
DB	1 GGGCAGCGGCTGCAGCGATTGCCATGCGCGCGCTGTGACTACGACGACAGCTGTG 60
OY	82 CTGGAATGCTCTGCACACTGACGGGCTAGTAGTGTGCCCCGGGGGCTGCGCGGATCGG 141
DB	61 CTGGAATGCTCTGCACACTGACGGGCTAGTAGTGTGCCCCGGGGGCTGCGCGGATCGG 120

QY	142	CTCCTCTACCACTTTTCTCCAGCTGACACTGGCCACCACAGCTGCTGGTGTGGTGTCTCAAC	201
Db	121	CTCCCTCTACCACTTTTCTCCAGCTGACACTGGCCACCACAGCTGCTGGTGTGGTGTCTCAAC	180
QY	202	ACGCAGCGCGCGGAGGAGGATTTTATCAATACACTGACGTGAACATACAGAGTTGAAAC	261
Db	181	ACGCAGCGCGCGGAGGAGGATTTTATCAATACACTGACGTGAACATACAGAGTTGAAAC	240
QY	262	CTCCCTGGCGCTGTACAAATGAATCACAAGCACAGCTGGCTATGAACTTACACACAA	321
Db	241	CTCCCTGGCGCTGTACAAATGAATCACAAGCACAGCTGGCTATGAACTTACACACAA	300
QY	322	GGTGGTTTATATTTGGGACAGTAGATCACTTGGTGGTACATCTTGACATATGAATA	381
Db	301	GGTGGTTTATATTTGGGACAGTAGATCACTTGGTGGTACATCTTGACATATGAATA	360
QY	382	CCCTTCAGATTTAATTTACTGGCATCTTGGTGTATAGACCCACAGATTAATCGATCTTGT	441
Db	361	CCCTTCAGATTTAATTTACTGGCATCTTGGTGTATAGACCCACAGATTAATCGATCTTGT	420
QY	442	CAGAAGACATTCATCTTGGCGCTCTTTCGCCAGAAAAACAAGCTGTTTATTTAAAGT	501
Db	421	CAGAAGACATTCATCTTGGCGCTCTTTCGCCAGAAAAACAAGCTGTTTATTTAAAGT	480
QY	502	TTACAGACAAATGCTGTGTGGCTTTGATGACTGTGTTTGTTCATGTGCAAAAGACTGATGAGA	561
Db	481	TTACAGACAAATGCTGTGTGGCTTTGATGACTGTGTTTGTTCATGTGCAAAAGACTGATGAGA	540
QY	562	AATCTTTTGTGAGAAACGTATATCTGTGCCAAGTTGCCATGTATCAAGTAAATCATATT	621
Db	541	AATCTTTTGTGAGAAACGTATATCTGTGCCAAGTTGCCATGTATCAAGTAAATCATATT	600
QY	622	TTAGAAACAGCAAAACCGTAAGTTGTGAATAATCCATGTTTCTGTGACCTACCATAGCTT	681
Db	601	TTAGAAACAGCAAAACCGTAAGTTGTGAATAATCCATGTTTCTGTGTACACTACCATAGCTT	660
QY	682	GCTATACAGACTGCTATATCTGAGCAATTTTAAATGATGTCTAAAGCACTAAATGCCAT	741
Db	661	GCTATACAGACTGCTATATCTGAGCAATTTTAAATGATGTCTAAAGCACTAAATGCCAT	720
QY	742	AACCCATGCGTTGAAGTGGAAAGATTTATCTTTGAAAAATGCTATTTGAAAAACCTTTTGAC	801
Db	721	AACCCATGCGTTGAAGTGGAAAGATTTATCTTTGAAAAATGCTATTTGAAAAACCTTTTGAC	780
QY	802	AAGACAATCCGCATTTATCTGGAATCTCTTGTGGCACCAAGCTTGAGCAAGCTAAATCC	861
Db	781	AAGACAATCCGCATTTATCTGGAATCTCTTGTGGCACCAAGCTTGAGCAAGCTAAATCC	840
QY	862	TTAGTTGAGAAATTTGAAGATTTTACGAACCTTGCTGAGATATCTCTCAGATATGATGT	921
Db	841	TTAGTTGAGAAATTTGAAGATTTTACGAACCTTGCTGAGATATCTCTCTAGATATGATGT	900
QY	922	GTCACATTTCTTAAATCTTCTGGAATCTCTGAGAGCAACGGAAAAAGCTTTTGGTCAGAT	981
Db	901	GTCACATTTCTTAAATCTTCTGGAATCTCTGAGAGCAACGGAAAAAGCTTTTGGTCAGAT	960
QY	982	TGAGTTGGCGTCTTCTGACTCAGACACCTCCATGTTTAAATGTCTGACACAAGGTT	1041
Db	961	TGAGTTGGCGTCTTCTGACTCAGACACCTCCATGTTTAAATGTCTGACACAAGGTT	1020
QY	1042	TATCATCTTCCAGATGCCAAATAGTATGAATAAAGAAAAAATATCTGAAAAAATGCAAAAT	1101
Db	1021	TATCATCTTCCAGATGCCAAATAGTATGAATAAAGAAAAAATATCTGAAAAAATGCAAAAT	1080
QY	1102	AAAGAAGGGAGAAACAACAAAAGGAAGCTGCTCTAGAAAGCAACCCAAAGTGGAGGCA	1161
Db	1081	AAAGAAGGGAGAAACAACAAAAGGAAGCTGCTCTAGAAAGCAACCCAAAGTGGAGGCA	1140
QY	1162	CTGACTGAAGTATTAAGAAATTTAGAGCAAGAAAAATTAAGAGAGTCAAGCTCTGGTGT	1221
Db	1141	CTGACTGAAGTATTAAGAAATTTAGAGCAAGAAAAATTAAGAGAGTCAAGCTCTGGTGT	1200
QY	1222	CCAGGTCAAGTACTGATTTGTGCAAGTATGACCGAAACATGTTCCAGCTGAGAGACTAT	1281

Db 1201 CCAGGTCAGTACTGATTTGTGCAAGTGAAGCCGAACTGTCACGCTGAGAGACTAT 1260
 QY 1282 ATCACTCTTGAGACGGAGCCCTCTTATTTAGAGCTCTACAGAAACCTTTGAGAAGAT 1341
 Db 1261 ATCACTCTTGAGACGGAGCCCTCTTATTTAGAGCTCTACAGAAACCTTTGAGAAGAT 1320
 QY 1342 AGCAAAAGCTGAAGAGTCTGATGAATTTAGAGAAAGACAGTTCAAGAGATTAAG 1401
 Db 1321 ACCAAAGCTGAAGAGTCTGATGAATTTAGAGAAAGACAGTTCAAGAGATTAAG 1380
 QY 1402 AATCTCACAAAAGACCTTAAGACCCCAAAACAAAGAGGGCTCTCCAAAGAAAGA 1461
 Db 1381 AATCTCACAAAAGACCTTAAGACCCCAAAACAAAGAGGGCTCTCCAAAGAAAGA 1440
 QY 1462 ACCCTCAAAAAGAAAAAGGAAAGTTGACCTTAACCTAAATGGTAGAAACCTGAGAA 1521
 Db 1441 ACCCTCAAAAAGAAAAAGGAAAGTTGACCTTAACCTAAATGGTAGAAACCTGAGAA 1500
 QY 1522 CTGGAAGAGGAAGAGATGTCGAGGAAGATATCTGAGAAATTAAGCAGTACCCAGAA 1581
 Db 1501 CTGGAAGAGGAAGAGATGTCGAGGAAGATATCTGAGAAATTAAGCAGTACCCAGAA 1560
 QY 1582 AGCTGCCGGAAGAAATTAAGCATGAAGATTTGATGAATTTGTCATCGGATGCTGCT 1641
 Db 1561 AGCTGCCGGAAGAAATTAAGCATGAAGATTTGATGAATTTGTCATCGGATGCTGCT 1620
 QY 1642 TTGGAATCTGTAAGAAAGAACCCCTCACTATCATCCGCTTCTGGGTTGACAGCAGCC 1701
 Db 1621 TTGGAATCTGTAAGAAAGAACCCCTCACTATCATCCGCTTCTGGGTTGACAGCAGCC 1680
 QY 1702 TATGCTCTGACAAAGGTACTACATGAAGTGAAGCCAAAGATAGCTGGTTCTTAATGACGA 1761
 Db 1681 TATGCTCTGACAAAGGTACTACATGAAGTGAAGCCAAAGATAGCTGGTTCTTAATGACGA 1740
 QY 1762 GAGCTAACCTTTGTCGAGCTGGAATTTACAGGGCCAGTAGGCTTGGGAAACCTCTG 1821
 Db 1741 GAGCTAACCTTTGTCGAGCTGGAATTTACAGGGCCAGTAGGCTTGGGAAACCTCTG 1800
 QY 1822 AAGGTTTACTTTCTTATATACGAGGTTCAACTGAGAAACAAGCTATCTCACTGCTTTG 1881
 Db 1801 AAGGTTTACTTTCTTATATACGAGGTTCAACTGAGAAACAAGCTATCTCACTGCTTTG 1860
 QY 1882 CGGAAAGAAAGAGAGCTTTTGAAAACTCATAGGGAAAAAGACAGCATGTTGTCCT 1941
 Db 1861 CGGAAAGAAAGAGAGCTTTTGAAAACTCATAGGGAAAAAGACAGCATGTTGTCCT 1920
 QY 1942 GAAGAAAGAGAGAGAGATGAANCAACTTAGACCTAGTAAGAGGACAGCATCTGCA 2001
 Db 1921 GAAGAAAGAGAGAGAGATGAANCAACTTAGACCTAGTAAGAGGACAGCATCTGCA 1980
 QY 2002 GATGTTTCCACTGACACTGGAAGCCGGTGGCCAGAAACAAGATGTCACAGCAAAAG 2061
 Db 1981 GATGTTTCCACTGACACTGGAAGCCGGTGGCCAGAAACAAGATGTCACAGCAAAAG 2040
 QY 2062 ATAGTTTGATATGCGTAATTTGCAAGTGAAGCTTCATCTCTGATCCATGCTCGGGG 2121
 Db 2041 ATAGTTTGATATGCGTAATTTGCAAGTGAAGCTTCATCTCTGATCCATGCTCGGGG 2100
 QY 2122 ATTGACATTGAACCCGTGACTTTAGAGTTGAGATTACATCTCTCAGCAAAATGTGC 2181
 Db 2101 ATTGACATTGAACCCGTGACTTTAGAGTTGAGATTACATCTCTCAGCAAAATGTGC 2160
 QY 2182 GTGAGAGCGAAGATATCAGTATTTAATCGGCTCTTTAAATACGGCCGCTCTACAG 2241
 Db 2161 GTGAGAGCGAAGATATCAGTATTTAATCGGCTCTTTAAATACGGCCGCTCTACAG 2220
 QY 2242 CAGTGATCTCATATGTCGGCTACTACAAGCGTCCCGTCTCTGATTGAGTTGACCT 2301
 Db 2221 CAGTGATCTCATATGTCGGCTACTACAAGCGTCCCGTCTCTGATTGAGTTGACCT 2280
 QY 2302 AGCAAGCCTTCTCTCTACTCCGAGGTGCTTCTCAGAGATCTCCAGCATATGAC 2361

Db 2281 AGCAAGCCTTCTCTCTACTCTCCAGGTGCCCTTTGACAGAGATCTCCAGCAATGAC 2340
 QY 2362 ATAGTTTCCAAATCAGTCTTCTTACACTTCACTTCCAGACTAGAGATTCGTGTC 2421
 Db 2341 ATAGTTTCCAAATCAGTCTTCTTACACTTCACTTCCAGACTAGAGATTCGTGTC 2400
 QY 2422 CCCTCTCTCATCAAGCGCGGAGTTGTTAGAGAGCTGAATCAAGCAAGCCACAGCT 2481
 Db 2401 CCCTCTCTCATCAAGCGCGGAGTTGTTAGAGAGCTGAATCAAGCAAGCCACAGCT 2460
 QY 2482 GATGGGCGACAGCACTGGCCATTACAGAGANTCTGAANCCCTTCGAGTCAGAGAG 2541
 Db 2461 GATGGGCGACAGCACTGGCCATTACAGAGANTCTGAANCCCTTCGAGTCAGAGAG 2520
 QY 2542 TATATCTCTGTCGCCCAAGCTCTTGTAAAAATGCCAGGGGTGAATGCCAAAACTGC 2601
 Db 2521 TATATCTCTGTCGCCCAAGCTCTTGTAAAAATGCCAGGGGTGAATGCCAAAACTGC 2580
 QY 2602 CGCTCTTATGACACAGCTTAAAGAACATGCGAGAAATGACAGCCCTGTACAGAGAG 2661
 Db 2581 CGCTCTTATGACACAGCTTAAAGAACATGCGAGAAATGACAGCCCTGTACAGAGAG 2640
 QY 2662 CTCACAGATATCTGGGGAAATGCTGCAATGCCAAACAGCTTTATGATTTCAATTCACAC 2721
 Db 2641 CTCACAGATATCTGGGGAAATGCTGCAATGCCAAACAGCTTTATGATTTCAATTCACAC 2700
 QY 2722 TCTTTGCAAGAGTCTATCAAAAAGGAAAAAGTGAACAGTGAATGGCTGTTTC 2781
 Db 2701 TCTTTGCAAGAGTCTATCAAAAAGGAAAAAGTGAACAGTGAATGGCTGTTTC 2760
 QY 2782 TATCCAGTCCGTACTTTTCAAGCGCTCTTGCAGACATCATAGTCAATTAATTT 2841
 Db 2761 TATCCAGTCCGTACTTTTCAAGCGCTCTTGCAGACATCATAGTCAATTAATTT 2820
 QY 2842 ATTGCTTCTATTTATCTTTTCCATGCTTTATATATTTGACGGTACAGCAAGC 2901
 Db 2821 ATTGCTTCTATTTATCTTTTCCATGCTTTATATATTTGACGGTACAGCAAGC 2880
 QY 2902 CAG 2904
 Db 2881 CAG 2883

RESULT 3
 AAF16235
 ID AAF16235 standard: cDNA: 2900 BP.
 XX
 AC AAF16235;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:670.
 XX
 OS Homo sapiens.
 PN WO20005174-A1.
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX

Human; prostate cancer; prostate cancer antigen: detection; diagnosis;
 neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
 vulnery; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;
 antibacterial; gene therapy; neural; immune; reproductive; renal;
 gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 wound; infectious disease; ss.

PI Rosen CA, Ruben SM:
XX WPI: 2000-587513/55.
DR P-PSDB; AAB57032.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 1; Page 1106-1107; 2338bp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotoxic, antiinfective, gynecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15566 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 2900 BP; 862 A; 635 C; 670 G; 728 T; 5 other;
Query Match 74.1%; Score 2857.4; DB 21; Length 2900;
Best Local Similarity 99.6%; Pred. NO. 0;
Matches 2865; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 25 CAGCGGCTGCAGCAGATGTCAGTGGCCGCGCTGCTGAGTACGAGCAGACACTGCTGTG 84
Db 15 CGCGCGGTGAGAGATGTCAGTGGCCGCGCTGCTGAGTACGAGCAGCA -CTGTGTGTG 73
QY -85 GAATGCTCCAGCATGACGGGCTAGTGTGTGCGCGCGCGGCTGCGCGGACCGCTC 144
Db 74 GAATGCTCCAGCATGACGGGCTAGTGTGTGCGCGCGCGGCTGCGCGGACCGCTC 133
QY 145 CTACACACTTTCACAGCTCAGTGGCCAGCCAGCTGCGTGGTGGTGGTCTCAACAG 204
Db 134 CTACACACTTTCACAGCTCAGTGGCCAGCCAGCTGCGTGGTGGTGGTCTCAACAG 193
QY 205 CAGCGGCGGAGAGAGATTTTATCAATCAGCTGAAGATAGAAGAGTGAACACCTC 264
Db 194 CAGCGGCGGAGAGAGATTTTATCAATCAGCTGAAGATAGAAGAGTGAACACCTC 253
QY 265 CCTGCGGCTGTACAAATGAATCACAAGCAGTGCCTATGAAGTTTACACAAAGT 324
Db 254 CCTGCGGCTGTACAAATGAATCACAAGCAGTGCCTATGAAGTTTACACAAAGT 313
QY 325 GGTGTATATTTGGCAGAGATGATCTGTGTGCTTGTGCTGACCTGACGATAGAACTCCT 384
Db 314 GGTGTATATTTGGCAGAGATGATCTGTGTGCTTGTGCTGACCTGACGATAGAACTCCT 373
QY 385 TCAGATTTAATTAATGATGATCTGTGTATAGAGCCACAGAAATATGATCTGTGCAA 444
Db 374 TCAGATTTAATTAATGATGATCTGTGTATAGAGCCACAGAAATATGATCTGTGCAA 433
QY 445 GAAGCATTCATCTTGGCGCTCTTTCGCCAGAAAAAACAAGCTGTTTATTAAAGCTTTC 504
Db 434 GAAGCATTCATCTTGGCGCTCTTTCGCCAGAAAAAACAAGCTGTTTATTAAAGCTTTC 493
QY 505 ACAGACATGCTGTGCTTGTGCTTGTGATCTGTGTGCTGATGGAAGAGTATAGAAAT 564
Db 494 ACAGACATGCTGTGCTTGTGCTTGTGATCTGTGTGCTGATGGAAGAGTATAGAAAT 553
QY 565 CTTTGTGAGAGAACTGATCTGTGCGAAGTTCCATGAGAGATGAATCTATTTTAA 624
Db 554 CTTTGTGAGAGAACTGATCTGTGCGAAGTTCCATGAGAGATGAATCTATTTTAA 613
QY 625 GAACAGCAAAACCTGAAGTTGTAGAAATCCATGTTCTATGACACCTACCATGCTTGTCT 684

Db 614 GAACAGCAAAACCTGAAGTTGTAGAAATCCATGTTCTATGACACCTACCATGCTTGTCT 673
QY 685 ATACAGACTGTATACAGTACGACATTTTAAATGCAATGCTTAAAGGAACTAAATGCGCATAC 744
Db 674 ATACAGACTGTATACAGTACGACATTTTAAATGCAATGCTTAAAGGAACTAAATGCGCATAC 733
QY 745 CCATGCGCTGAAGTGAAGATTTATCTTAAAGAAATGCTATTTGAAAAACCTTTTGACAAAG 804
Db 734 CCATGCGCTGAAGTGAAGATTTATCTTAAAGAAATGCTATTTGAAAAACCTTTTGACAAAG 793
QY 805 ACAATCCGCCATTAATCTGATCTTGTGCGCAGCTTGGAGCCAAAGACTAAATCCTTA 864
Db 794 ACAATCCGCCATTAATCTGATCTTGTGCGCAGCTTGGAGCCAAAGACTAAATCCTTA 853
QY 865 GTTAGAGATTTGAAGATTTAGCAACTTGTGAGATCTCTCTGATGATGATGTGTC 924
Db 854 GTTAGAGATTTGAAGATTTAGCAACTTGTGAGATCTCTCTGATGATGATGTGTC 913
QY 925 ACATTTCTTAATCTTCTGAAATCTCTGAGCAACGGAAGAAAGCTTTGGTCAGAAATYCA 984
Db 914 ACATTTCTTAATCTTCTGAAATCTCTGAGCAACGGAAGAAAGCTTTGGTCAGAAATYCA 973
QY 985 GGTGCGCTGTTCTTGTGCTGACCTCCATCTGTTTATTAATGCTCGAGCAAGGTTTAT 1044
Db 974 GGTGCGCTGTTCTTGTGCTGACCTCCATCTGATTTTAAATGCTCGAGCAAGGTTTAT 1033
QY 1045 CATCTCCAGATGCGCAAAATAGTAAAAAACAAGAAATATCTGAAAAATGGAATTTAA 1104
Db 1034 CATCTCCAGATGCGCAAAATAGTAAAAAACAAGAAATATCTGAAAAATGGAATTTAA 1093
QY 1105 GAAGGGGAAGAAACAAAAAGAACTGCTCTAGAAAGCAACCCAAAGTGGAGGCACTG 1164
Db 1094 GAAGGGGAAGAAACAAAAAGAACTGCTCTAGAAAGCAACCCAAAGTGGAGGCACTG 1153
QY 1165 ACTGAAGTATTAAGAAATTTGAGGCAAGAAATTAAGAGATGAAAGCTTGTGGTGTCCA 1224
Db 1154 ACTGAAGTATTAAGAAATTTGAGGCAAGAAATTAAGAGATGAAAGCTTGTGGTGTCCA 1213
QY 1225 GGTCAAGTACTGATTTGTGCAAGTATGACCAAGTATGTTCCAGTCCAGAGACTATATC 1284
Db 1214 GGTCAAGTACTGATTTGTGCAAGTATGACCAAGTATGTTCCAGTCCAGAGACTATATC 1273
QY 1285 ACTCTTGAGCGGAGGCGCTTCTTATTAAGGCTCTACAGAAACCTTTGAGAGGATAGC 1344
Db 1274 ACTCTTGAGCGGAGGCGCTTCTTATTAAGGCTCTACAGAAACCTTTGAGAGGATAGC 1333
QY 1345 AAAGCTGAAGAGTCTGATGATTAATTTAGAGAAAGACAGTTCAAGAGATTAAGGAAA 1404
Db 1334 AAAGCTGAAGAGTCTGATGATTAATTTAGAGAAAGACAGTTCAAGAGATTAAGGAAA 1393
QY 1405 TCTCAGAAAAGACCTAAAGACCCCAAAACAAAGAGGCGCTTCTACCAAGAAAGAAC 1464
Db 1394 TCTCAGAAAAGACCTAAAGACCCCAAAACAAAGAGGCGCTTCTACCAAGAAAGAAC 1453
QY 1465 CTCAAAAAAGAAAAAGGAACTTGAACCTTAACCTAAATGTTAGGAAAACTGAAAGAACTG 1524
Db 1454 CTCAAAAAAGAAAAAGGAACTTGAACCTTAACTCAATATGTTAGGAAAACTGAAAGAACTG 1513
QY 1525 GAAGAGAGAGAGATGTCGAGGAAAGATATGCTGAGAAATTAACAGTACGCCAGAAAGC 1584
Db 1514 GAAGAGAGAGAGATGTCGAGGAAAGATATGCTGAGAAATTAACAGTACGCCAGAAAGC 1573
QY 1585 TGCCCGGAGAAATTAAGCATGAAGATTTGATGAAATTTGTCATGCGATGCTGCTTTC 1644
Db 1574 TGCCCGGAGAAATTAAGCATGAAGATTTGATGAAATTTGTCATGCGATGCTGCTTTC 1633
QY 1645 GGAATCTGAAGAAACCCCTCACTATCATCATCCGCTTCTGCGTGGAGGAGCCACCTAT 1704
Db 1634 GGAATCTGAAGAAACCCCTCACTATCATCATCCGCTTCTGCGTGGAGGAGCCACCTAT 1693
QY 1705 GCTCTGACAAAGGTAATCAATGAAGTGAAGCAAGATACGTTCTTATGACGCAAG 1764

```

Db 1694 GCTGTGACAAAGGTACTACATGAGTGAAGCCAGATACGTTCTTTATGACGACAG 1753
Qy 1765 CTACCTTTGTTGGCAGCTGTAATTTACAGGCGAGTAGGCGTGGGAAACCTGTGAG 1824
Db 1754 CTACCTTTGTTGGCAGCTGTAATTTACAGGCGAGTAGGCGTGGGAAACCTGTGAG 1813
Qy 1825 GTTACTTTCTTATATAGGAGGTTCAACCTGAGGAACAAGCTATCTCAGCTTTGCGG 1884
Db 1814 GTTACTTTCTTATATAGGAGGTTCAACCTGAGGAACAAGCTATCTCAGCTTTGCGG 1873
Qy 1885 AAAGAAAAGAAAGCTTTGAAAAAACTCATAGGAAAAAGCAAGATGGTTCCCTGAA 1944
Db 1874 AAAGAAAAGAAAGCTTTGAAAAAACTCATAGGAAAAAGCAAGATGGTTCCCTGAA 1933
Qy 1945 GAAAGAGAAGGAGATGAACAACTTAGACCTAGTAAGAGGACACAGATCTGCAGAT 2004
Db 1934 GAAAGAGAAGGAGATGAACAACTTAGACCTAGTAAGAGGACACAGATCTGCAGAT 1993
Qy 2005 GTTTCACACTGACACTCGGAAAGCCGCTGGCCAGAAACAGATGTAACAGCAAGACATA 2064
Db 1994 GTTTCACACTGACACTCGGAAAGCCGCTGGCCAGAAACAGATGTAACAGCAAGACATA 2053
Qy 2065 GTTGTGATATGCGTGAATTTGAAAGTGAGCTTGCATCTGTATCCATCGTGGGGCAT 2124
Db 2054 GTTGTGATATGCGTGAATTTGAAAGTGAGCTTGCATCTGTATCCATCGTGGGGCAT 2113
Qy 2125 GACATTGAACCCGCTGACTTTAAGTTGAGATTAATCTCTACTCCAGAAATGTGCGTG 2184
Db 2114 GACATTGAACCCGCTGACTTTAAGTTGAGATTAATCTCTACTCCAGAAATGTGCGTG 2173
Qy 2185 GAGCCCAAGAGATATGATTTAATCGCTTTAAATAAGCGCCGCTCTACAGCCAG 2244
Db 2174 GAGCCCAAGAGATATGATTTAATCGCTTTAAATAAGCGCCGCTCTACAGCCAG 2233
Qy 2245 TGATCTCATATCTCCCGCTACTACAAAGCGTCCGCTGCTTGATTTGAGTTGACCTTAC 2304
Db 2234 TGATCTCATATCTCCCGCTACTACAAAGCGTCCGCTGCTTGATTTGAGTTGACCTTAC 2293
Qy 2305 AAGCCTTTCTCTCTACTTCCCGAGTGGCTGTTCCAGGAGATCTCCAGCAATGACAT 2364
Db 2294 AAGCCTTTCTCTCTACTTCCCGAGTGGCTGTTCCAGGAGATCTCCAGCAATGACAT 2353
Qy 2365 AGTTCCAAACTCAGCTCTTTTACACTTCCCGAGACTAGGATTTCTGTGCCCC 2424
Db 2354 AGTTCCAAACTCAGCTCTTTTACACTTCCCGAGACTAGGATTTCTGTGCCCC 2413
Qy 2425 TCTCTCATGCAAGCGCGGAGTTGTTGAGAGCTGAAACAAGCAAGCCACAGCTGAT 2484
Db 2414 TCTCTCATGCAAGCGCGGAGTTGTTGAGAGCTGAAACAAGCAAGCCACAGCTGAT 2473
Qy 2485 GCGGCGAAGACAGCTGCGCATTTACAGCAGATTTCCGAGTCCAGAGAGAT 2544
Db 2474 GCGGCGAAGACAGCTGCGCATTTACAGCAGATTTCCGAGTCCAGAGAGAT 2533
Qy 2545 AATCTGTGCTCCCAAGACTTTCTTTAAAAATGCCAGGGGTGAATGCCAAAAATGCGCGC 2604
Db 2534 AATCTGTGCTCCCAAGACTTTCTTTAAAAATGCCAGGGGTGAATGCCAAAAATGCGCGC 2593
Qy 2605 TCTTGTATGACACACAGCTTAAGAACATCGCAGATTTACAGCCCTTCACAGACAGCTC 2664
Db 2594 TCTTGTATGACACACAGCTTAAGAACATCGCAGATTTACAGCCCTTCACAGACAGCTC 2653
Qy 2665 ACGATATATTCGGGGAATGCTGCAAAATGCCAAACAGCTTTATGATTTCACTACACCTCT 2724
Db 2654 ACGATATATTCGGGGAATGCTGCAAAATGCCAAACAGCTTTATGATTTCACTACACCTCT 2713
Qy 2725 TTTGCGAAGTGTATCAAAAAGAAAAAGAAAAAGTAAACAGTATGCTGTTTTCTTA 2784
Db 2714 TTTGCGAAGTGTATCAAAAAGAAAAAGAAAAAGTAAACAGTATGCTGTTTTCTTA 2773
Qy 2785 TCCCATGCTCTACTTTTCAGCGGCTCTTGCCAGACATCATAGTCAATTATTAATTAT 2844
Db 2774 TCCCATGCTCTACTTTTCAGCGGCTCTTGCCAGACATCATAGTCAATTATTAATTAT 2833

```

```

Qy 2845 GGTTCGATATTTCATTCTTTTCCAAATGCTTAATGATTGTACGGTGSACAGAGCCAG 2904
Db 2834 GGTTCGATATTTCATTCTTTTCCAAATGCTTAATGATTGTACGGTGSACAGAGTTCTAG 2893

RESULT 4
AAS63230
ID AAS63230 standard; cDNA; 2718 BP.
XX
AC AAS63230;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human full-length XPF (ERCC4) DNA.
XX
KW DNA mutation-binding protein; nuclease; DNA mismatch; cancer; PCR primer;
KW DNA damage; human xeroderma pigmentosum complementation group; XPF; XPA;
KW XPC; XPE; ERCC4; human Muts homologue 2; hMSH2; Muts; Nuc; MutY; Fpg; ss;
KW Fapy-DNA glycosylase; uracil DNA glycosylase; ung; Tdg; xthA gene; Uvr A;
KW A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;
KW thymine DNA-glycosylase; Uvr B; Uvr C; nth gene; nfo gene; exonuclease;
KW endonuclease.
XX
OS Homo sapiens.
XX
PN WO200173079-A2.
XX
PD 04-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-US09700.
XX
PR 28-MAR-2000; 2000US-192764P.
XX
PR 29-AUG-2000; 2000US-0650855.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Mc Cutchen-maloney SL;
XX
DR WPI: 2001-656920/75.
XX
DR P-PSDB: AA069743.
XX
PT Recombinant chimeric protein, useful for detecting and quantifying DNA
PT mutations, e.g. in disease diagnosis, comprises mutation-binding
PT protein and nuclease
XX
PS Disclosure; Page 69-70; 128p; English.
XX
CC Sequences AAS63224-AAS63305 represent cDNA encoding and PCR primers used
CC to amplify cDNA encoding proteins which can be used in the synthesis of
CC chimeric proteins comprising a DNA mutation-binding protein, a linker and
CC a nuclease, by recombinant technology. The chimeric proteins are useful
CC for detection, quantification and mapping of DNA sequence variations
CC including mutations, for example, caused by damage and mismatches. The
CC proteins are able to bind to the site of the DNA mutation and cut it out
CC of the molecule. This is useful for early diagnosis of cancer and other
CC diseases. The proteins used in the invention include human XPF (or
CC ERCC4), human xeroderma pigmentosum complementation groups A, C and E
CC (XPA, XPC and XPE), human Muts homologue 2 (hMSH2), Serratia marcescens
CC nuclease (Nuc), Thermus thermophilus Muts, Escherichia coli Fapy-DNA
CC glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific adenine
CC glycosylase (MutY), synthetic T4 endonuclease V (T4 endo V), thymine
CC DNA-glycosylase (Tdg), E. coli Uvr A, B and C, and E. coli endonucleases
CC and exonucleases.
XX
SO Sequence 2718 BP; 827 A; 588 C; 634 G; 669 T; 0 other;

Query Match 70.5%; Score 2718; DB 22; Length 2718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 ATGGCGCGCTGCTGAGTACAGCAGCAGCTGTGTGGAAGTCTCTGACACTGACGG 105
|||||

```

Db 1 ATGGCCCGCTGCTGAGTACGACGACAGCTGTGCTGGAAGCTGCTGACACTGACGCG 60
QY 106 CTAGTAGTGTGCG 165
Db 61 CTGATGTGTGCG 120
QY 166 CACTGGCACCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
Db 121 CACTGGCACCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 226 TTTATCATCATGCTGAAGATAGAAGAGAGTGAACACCTCCCTGCGCGCTGTAAACAAATGAA 285
Db 181 TTTATCATCATGCTGAAGATAGAAGAGTGAACACCTCCCTGCGCGCTGTAAACAAATGAA 240
QY 286 ATACAAAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
Db 241 ATACAAAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 346 AGGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 405
Db 301 AGGATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 406 TTGGTGTATGAGCCCGACAAATATGAGTCTTGTCAAGAGCATTCATCTTGGCCTC 465
Db 361 TTGGTGTATGAGCCCGACAAATATGAGTCTTGTCAAGAGCATTCATCTTGGCCTC 420
QY 466 TTTCGGCAGAAAAACAAGCTGTTTATTAAGCTTTCACAGCATGCTTGGCTT 525
Db 421 TTTCGGCAGAAAAACAAGCTGTTTATTAAGCTTTCACAGCATGCTTGGCTT 480
QY 526 GATACCTGTTTTGTCATGTGAGAAAGATGAGAAATCTTTTGTGAGAAACCTGAT 585
Db 481 GATACCTGTTTTGTCATGTGAGAAAGATGAGAAATCTTTTGTGAGAAACCTGAT 540
QY 586 CTGTGGCCAGGTTCCATGTAGCAGTAACTCATTTTAAACAGCAACCAACCTGAAAGTT 645
Db 541 CTGTGGCCAGGTTCCATGTAGCAGTAACTCATTTTAAACAGCAACCAACCTGAAAGTT 600
QY 646 GTAGAAATCCATGTTTCTATGACACCTACCATGCTGCTTACAGATGCTATCTGGAC 705
Db 601 GTAGAAATCCATGTTTCTATGACACCTACCATGCTGCTTACAGATGCTATCTGGAC 660
QY 706 ATTTAAATCATGTCTAAAGAACTAAATGCCATTAACCATGCTTGAAGTGAAGAT 765
Db 661 ATTTAAATCATGTCTAAAGAACTAAATGCCATTAACCATGCTTGAAGTGAAGAT 720
QY 766 TTTATCTTAAAGAAATGCTATGAGAAACCTTGTGACAAAGCAATCCGCATATCTGAT 825
Db 721 TTTATCTTAAAGAAATGCTATGAGAAACCTTGTGACAAAGCAATCCGCATATCTGAT 780
QY 826 CTTTGTGGCAGCAGCTTGTGAGCCAAAGCTAAATCCTTAGTTCAGATTTGAAGATTTA 885
Db 781 CTTTGTGGCAGCAGCTTGTGAGCCAAAGCTAAATCCTTAGTTCAGATTTGAAGATTTA 840
QY 886 CGAATTTGCTGCAATGTCTCTCACTATGATGTGTCAATTTCTTAACTCTTCTGAA 945
Db 841 CGAATTTGCTGCAATGTCTCTCACTATGATGTGTCAATTTCTTAACTCTTCTGAA 900
QY 946 TCTCTGAGACCAAGGAAAAAGCTTTTGTGACAAATCAAGTTGGCTGTTTCTGACTCC 1005
Db 901 TCTCTGAGACCAAGGAAAAAGCTTTTGTGACAAATCAAGTTGGCTGTTTCTGACTCC 960
QY 1006 AGACACCTGATGTTTAAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAATG 1065
Db 961 AGACACCTGATGTTTAAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAATG 1020
QY 1066 AGTAAAAAAGAAAAATATCTGAAAAAATGAAATTAAGAAAGGGAAGAAACAAAAAG 1125
Db 1021 AGTAAAAAAGAAAAATATCTGAAAAAATGAAATTAAGAAAGGGAAGAAACAAAAAG 1080
QY 1126 GAAGTGTCTAGAAAGCAACCAAGTGGAGGCACTGACTGAAGTATTTAAAGAAATTT 1185
Db 1081 GAAGTGTCTAGAAAGCAACCAAGTGGAGGCACTGACTGAAGTATTTAAAGAAATTT 1140

QY 1186 GAGGCAAGAAATAGAGAGAGTGAAGCTTTGGTGTCCAGGTCAACTACATGATTTGTGCA 1245
Db 1141 GAGGCAAGAAATAGAGAGAGTGAAGCTTTGGTGTCCAGGTCAACTACATGATTTGTGCA 1200
QY 1246 AGTATGACCGAACAATGTTCCACAGCTAGAGACTATATCACTTGTGAGCGGAGCCCTTC 1305
Db 1201 AGTATGACCGAACAATGTTCCACAGCTAGAGACTATATCACTTGTGAGCGGAGCCCTTC 1260
QY 1306 TTTATGAGGCTCTACAGAAAGAACTTTGAGAAGATAGCAAAAGCTGAAGAAAGTGTGATG 1365
Db 1261 TTTATGAGGCTCTACAGAAAGAACTTTGAGAAGATAGCAAAAGCTGAAGAAAGTGTGATG 1320
QY 1366 AAATTTAGAAAGAAAGCACTTCAAAAGAAATTTAGAAATTTAGAAAGCTTAAAGAC 1425
Db 1321 AAATTTAGAAAGAAAGCACTTCAAAAGAAATTTAGAAATTTAGAAAGCTTAAAGAC 1380
QY 1426 CCCCCAAAACAAGAGCGGCTTCTACCAAAAGAAAGAACTTCAAAAGAAAGAAAGCGAAG 1485
Db 1381 CCCCCAAAACAAGAGCGGCTTCTACCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGAAG 1440
QY 1486 TTGACCTTAACCTCAATGATGAGAAACCTGAAGAACTGGAAGAGGAGGATGTGAG 1545
Db 1441 TTGACCTTAACCTCAATGATGAGAAACCTGAAGAACTGGAAGAGGAGGATGTGAG 1500
QY 1546 GAAAGATATGCTGAGAAATTAAGCAGTAGCCAGAAAGCTGCCCCGGAAGAAATTAACAT 1605
Db 1501 GAAAGATATGCTGAGAAATTAAGCAGTAGCCAGAAAGCTGCCCCGGAAGAAATTAACAT 1560
QY 1606 GAAAGATTTGATGATTAATTTGTCACTGCGATGCTGCTTGGAAATCTTGAAGAAACCCCTC 1665
Db 1561 GAAAGATTTGATGATTAATTTGTCACTGCGATGCTGCTTGGAAATCTTGAAGAAACCCCTC 1620
QY 1666 ACTATCATCATGCGCTTGGGTTGAGAGGACCCCTATGCTGAGCAAGGATCTCAT 1725
Db 1621 ACTATCATCATGCGCTTGGGTTGAGAGGACCCCTATGCTGAGCAAGGATCTCAT 1680
QY 1726 GAAATGAGGCAAGATACGTGTTCTTTATGACAGCAAGCTTAACCTTGTGTTGGCAGCTT 1785
Db 1681 GAAATGAGGCAAGATACGTGTTCTTTATGACAGCAAGCTTAACCTTGTGTTGGCAGCTT 1740
QY 1786 GAAATTTACAGGCGAGTAGGCGCTGGGAAACCTGTAGAGGTTTTACTTTCTTATATACGA 1845
Db 1741 GAAATTTACAGGCGAGTAGGCGCTGGGAAACCTGTAGAGGTTTTACTTTCTTATATACGA 1800
QY 1846 GGTTCACATGAGGAAACGATCTCATGCTTGTGAGAAAGAAAGAAAGCTTTGAA 1905
Db 1801 GGTTCACATGAGGAAACGATCTCATGCTTGTGAGAAAGAAAGAAAGCTTTGAA 1860
QY 1906 AAATCTATTAAGGGAAGAAAGCAAGCATGTTTCCCTGAGAAAGAGAGCAGAGATGAA 1965
Db 1861 AAATCTATTAAGGGAAGAAAGCAAGCATGTTTCCCTGAGAAAGAGAGCAGAGATGAA 1920
QY 1966 ACAAACTTAGACCTAGTAAAGGACAGCATCTGCAAGTGTTCCTACTGACACTCGGAAA 2025
Db 1921 ACAAACTTAGACCTAGTAAAGGACAGCATCTGCAAGTGTTCCTACTGACACTCGGAAA 1980
QY 2026 GCGGTGGCCAGGAAGCAAGATGTTACAGCAAGCAAGTGTGATGATGATGATGATGATGAT 2085
Db 1981 GCGGTGGCCAGGAAGCAAGATGTTACAGCAAGCAAGTGTGATGATGATGATGATGATGAT 2040
QY 2086 CGAAGTGAAGTCTCATCTGATCATCTGATGATGATGATGATGATGATGATGATGATGAT 2145
Db 2041 CGAAGTGAAGTCTCATCTGATCATCTGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2146 GAGGTGAGATTAATCATCTCTCATCTCAGAAATGTGCTGAGCGCAAGAGATATCAGTAT 2205
Db 2101 GAGGTGAGATTAATCATCTCTCATCTCAGAAATGTGCTGAGCGCAAGAGATATCAGTAT 2160
QY 2206 TTTATGAGGCTTTAAATACGCGCGCTGCTACAGCGCAGTGCATCTCATGATGATGATGAT 2265
Db 2161 TTTATGAGGCTTTAAATACGCGCGCTGCTACAGCGCAGTGCATCTCATGATGATGATGATGAT 2220

Db 30242 GAAGTGTATCAAAAGAAAAGAAAAGTGAACAGTGCCTGTTTCTTATCCAT 30301
QY 2791 GCCTGACTTTTCACGGGCTCCTTGCAGACATCATAGTCTATATATTTGGTTG 2850
Db 30302 GCCTGACTTTTCACGGGCTCCTTGCAGACATCATAGTCTATATATTTGGTTG 30361
QY 2851 CTAATTCATCTTTTCCAAATGCTTTAATGATGTACGGTGACCAAGAGCAGATTCC 2910
Db 30362 CTAATTCATCTTTTCCAAATGCTTTAATGATGTACGGTGACCAAGAGCAGATTCC 30421
QY 2911 TCTGTAACTGTGAGTAGGACATCATTTGACCTGCTGCTCTTTTCTCTCC 2970
Db 30422 TCTGTAACTGTGAGTAGGACATCATTTGACCTGCTGCTCTTTTCTCTCC 30481
QY 2971 TGCACCGCTATAGCCGGGCTAGCATGTTCTTTTAATGAGGTTTGCAGATCAGGT 3030
Db 30482 TGCACCGCTATAGCCGGGCTAGCATGTTCTTTTAATGAGGTTTGCAGATCAGGT 30541
QY 3031 AAAGTTCTACAGATGATACAGAAAGTGAACCTTTACCTATCCTAAGATCTATT 3090
Db 30542 AAAGTTCTACAGATGATACAGAAAGTGAACCTTTACCTATCCTAAGATCTATT 30601
QY 3091 TAGAAAGAAATATGCTAAAGCTTGATGAGCGGTGACGGGAGGAAAAGACAGCACA 3150
Db 30602 TAGAAAGAAATATGCTAAAGCTTGATGAGCGGTGACGGGAGGAAAAGACAGCACA 30661
QY 3151 GAAAGTACCATTTTAAACAGTCTGTATCTAGTGAACAATTAATACATCTTAAT 3210
Db 30662 GAAAGTACCATTTTAAACAGTCTGTATCTAGTGAACAATTAATACATCTTAAT 30721
QY 3211 GCACCTATACCATGCTGCTGCTGCTCCAAATCTGGTCTTGTCTGTCTGTGA 3270
Db 30722 GCACCTATACCATGCTGCTGCTGCTCCAAATCTGTCTTGTCTGTCTGTGA 30781
QY 3271 CGCTTGAACTGATGTTTGTAGGAATCATGTTGACCCCTTGTCTAAGAGACCT 3330
Db 30782 CGCTTGAACTGATGTTTGTAGGAATCATGTTGACCCCTTGTCTAAGAGACCT 30841
QY 3331 TCTGGAACCTGAGAAAGAAACATCTTGGCAATCTGACAGATCTCTACACAT 3390
Db 30842 TCTGGAACCTGAGAAAGAAACATCTTGGCAATCTGACAGATCTCTACACAT 30901
QY 3391 TTTCTGAGCTCATACTTCTGCTGCTCTCTAAGAAATTCATGAGCCCTTCTACT 3450
Db 30902 TTTCTGAGCTCATACTTCTGCTGCTCTCTAAGAAATTCATGAGCCCTTCTACT 30961
QY 3451 ACTAATTCAGAGAGCTCTCTCAAAAACGTGGTGAATCTCTTAATGACCTTAACATA 3510
Db 30962 ACTAATTCAGAGAGCTCTCTCAAAAACGTGGTGAATCTCTTAATGACCTTAACATA 31021
QY 3511 TGTAGCATATACATAATTTTCAATTTGTTCCAAATTAATTTTAAAGCAAAATGAATTAC 3570
Db 31022 TGTAGCATATACATAATTTTCAATTTGTTCCAAATTAATTTTAAAGCAAAATGAATTAC 31081
QY 3571 CTGTTTGCAGAAAGTAAATGATGAGAGGCTTTAGAAATCTCAATTTTGCACATATCA 3630
Db 31082 CTGTTTGCAGAAAGTAAATGATGAGAGGCTTTAGAAATCTCAATTTTGCACATATCA 31141
QY 3631 GTTCTCAATATAGAGATCCCTAACTCAAGCTGGCTAGTTAAGAGTTTTCACACT 3690
Db 31142 GTTCTCAATATAGAGATCCCTAACTCAAGCTGGCTAGTTAAGAGTTTTCACACT 31201
QY 3691 CCGTGTTCACGCTTATATCTTAAGACACCATCATATCTCTAAGAAATTAACAC 3750
Db 31202 CCGTGTTCACGCTTATATCTTAAGACACCATCATATCTCTAAGAAATTAACAC 31261
QY 3751 TAATTCGAGTGAAGCCGAGATCGACACCTGACCCCTGCTGGGAGCAGAGTGAAGT 3810
Db 31262 TAATTCGAGTGAAGCCGAGATCGACACCTGACCCCTGCTGGGAGCAGAGTGAAGT 31321
QY 3811 TTGTCTCTATTAAGAAAGAAAGAAAGAAATTAACACCTTAACCTCA 3857
Db 31322 TTGTCTCTATTAAGAAAGAAAGAAAGAAATTAACACCTTAACCTCA 31368

RESULT 6
ABL64428
ID ABL64428 standard; DNA: 35641 BP.
XX
AC ABL64428;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2765.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 18-SEP-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237284P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;

DR WPI: 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1: SEQ ID 2765; 44p; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
CC to AB170110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC osteophagel, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 35641 BP; 10575 A; 7169 C; 7323 G; 10574 T; 0 other;
Query Match 47.4%; Score 1829.4; DB 24; Length 35641;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 2011 ACTGACACTCGGAAAGCGGTGGCCAGAGAGAAATGTGTACAGCAAGCATATGTTGG 2070
DB 29522 ACTTACTTTTCTCTGTGTAGTGGCCAGAGACAGATGTTACACAGCAAGCATATGTTG 29521
QY 2071 GATATGCGTGAATTTGGAAGTGAAGCTTCATCTCTGATCATCGTGGGGGATTGACATT 2130
DB 29582 GATATGCGTGAATTTGGAAGTGAAGCTTCATCTCTGATCATCGTGGGGGATTGACATT 29641
QY 2131 GAACCCGGAATTTAGAGTGGAGATTACATCTCTGATCATCGTGGGGGATTGACATT 2130
DB 29642 GAACCCGGAATTTAGAGTGGAGATTACATCTCTGATCATCGTGGGGGATTGACATT 29701
QY 2191 AAGATGATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2250
DB 29702 AAGATGATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 29761
QY 2251 TCCATGTCGCGCTACTACAAAGCGTCCGCTGCTGATTTGATTTGACCTAGCAAGCCT 2310
DB 29762 TCCATGTCGCGCTACTACAAAGCGTCCGCTGCTGATTTGATTTGACCTAGCAAGCCT 29821
QY 2311 TTCTCTCTCACTTCCCGAGGTGCTTGTTCAGAGATCTCCAGCAATGATTAATGTTCC 2370
DB 29822 TTCTCTCTCACTTCCCGAGGTGCTTGTTCAGAGATCTCCAGCAATGATTAATGTTCC 29881
QY 2371 AAATGCACTCTTTTACACTCTACCTCCCGAGCTAGAGATTTCTGTGTCGCCCTCTCT 2430
DB 29882 AAATGCACTCTTTTACACTCTACCTCCCGAGCTAGAGATTTCTGTGTCGCCCTCTCT 29941
QY 2431 CATGCAACGGGGAGTGTGTTGAGAGCTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2490
DB 29942 CATGCAACGGGGAGTGTGTTGAGAGCTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 30001
QY 2491 ACAGCACTGGGCAATTAAGCAGATTTGAAACCTTCCCGAGTCAAGAGATTAATATCT 2550
DB 30002 ACAGCACTGGGCAATTAAGCAGATTTGAAACCTTCCCGAGTCAAGAGATTAATATCT 30061
QY 2551 GGTCGCCAAGAGCTCTTTGTTAAATATGCGAGGGTGAATGCAAAATGCGCGCTCTCTTG 2610
DB 30062 GGTCGCCAAGAGCTCTTTGTTAAATATGCGAGGGTGAATGCAAAATGCGCGCTCTCTTG 30121
QY 2611 ATGCACCAAGTTAAGAACATCGCAGAAATTAGCAGCCTGTTCACAAAGAGAGACTCAGAGT 2670

DB 30122 ATGCACCAAGTTAAGAACATCGCAGAAATTAGACGCCCTGTACAAAGAGAGCTCAGAGT 30181
QY 2671 ATTCGGGGAATGCTGCAAAATGCCAAAGAGCTTTATGATTTTCAATTCACACCTTTTGCA 2730
DB 30182 ATTCGGGGAATGCTGCAAAATGCCAAAGAGCTTTATGATTTTCAATTCACACCTTTTGCA 30241
QY 2731 GAAGTCGATCAAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 2790
DB 30242 GAAGTCGATCAAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 30301
QY 2791 GCGTCTACTTTTCAAGCGGCTCTCTCCAGACATCATATGATTAATTTATTTGTTTG 2850
DB 30302 GCGTCTACTTTTCAAGCGGCTCTCTCCAGACATCATATGATTAATTTATTTGTTTG 30361
QY 2851 CTATTTCATCTTTTCCAAATGCTCTTAATGATTTAGCGGTGACCAAGCCAGGATTC 2910
DB 30362 CTATTTCATCTTTTCCAAATGCTCTTAATGATTTAGCGGTGACCAAGCCAGGATTC 30421
QY 2911 TCTGTGACCTGCGAGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2970
DB 30422 TCTGTGACCTGCGAGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 30481
QY 2971 TGACACGCTATGCGGGCTTAGCATGTTTCTTTTAAATGAGTTTGTCAAGATCAGT 3030
DB 30482 TGACACGCTATGCGGGCTTAGCATGTTTCTTTTAAATGAGTTTGTCAAGATCAGT 30541
QY 3031 AAAGTCTCTACAGATGATTTACAGAGATGAGAACTTACCTGATCTTACAGATTCATT 3090
DB 30542 AAAGTCTCTACAGATGATTTACAGAGATGAGAACTTACCTGATCTTACAGATTCATT 30601
QY 3091 TAGAAGGAATATGCTTAAGCCTGCGATGAGCGTCCAGAGAGGAAAGAGAGAGAGAGAG 3150
DB 30602 TAGAAGGAATATGCTTAAGCCTGCGATGAGCGTCCAGAGAGGAAAGAGAGAGAGAGAG 30661
QY 3151 GAAAGCTACATTTTAAACAGTCTTGTATCTAGTGAACATAAATTAACAGTCTTAAT 3210
DB 30662 GAAAGCTACATTTTAAACAGTCTTGTATCTAGTGAACATAAATTAACAGTCTTAAT 30721
QY 3211 GCACCTTATACCATGCTCTGTCGCTGCTCCAAATCTGCTGCTGCTGCTGCTGCTGCTG 3270
DB 30722 GCACCTTATACCATGCTCTGTCGCTGCTCCAAATCTGCTGCTGCTGCTGCTGCTGCTG 30781
QY 3271 CGCTGAACTGATGTTTGTGAGAAATCATGTTTGTGACCCCTTTGTCTACAAAGAGCCT 3330
DB 30782 CGCTGAACTGATGTTTGTGAGAAATCATGTTTGTGACCCCTTTGTCTACAAAGAGCCT 30841
QY 3331 TCTGGAACACTGAGAAAGAAATCATCTTGTGCAATTCCTGACAGTTCTCTACACATT 3390
DB 30842 TCTGGAACACTGAGAAAGAAATCATCTTGTGCAATTCCTGACAGTTCTCTACACATT 30901
QY 3391 TTCTTCACTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3450
DB 30902 TTCTTCACTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30961
QY 3451 ACTAATTAAGAGAGTCTCTCAAAAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3510
DB 30962 ACTAATTAAGAGAGTCTCTCAAAAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 31021
QY 3511 TGTAGCATATATATATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3570
DB 31022 TGTAGCATATATATATTTTCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 31081
QY 3571 CTGTTTGAAGAAATTAATGATGAGAGAGCTCTTGAATTTGATTTGATTTGATTTGATTT 3630
DB 31082 CTGTTTGAAGAAATTAATGATGAGAGAGCTCTTGAATTTGATTTGATTTGATTTGATTT 31141
QY 3631 GTCCTCTATATACAGAGATTCCTTAAGTCAAGTCTGATGATGATGATGATGATGATGATG 3690
DB 31142 GTCCTCTATATACAGAGATTCCTTAAGTCAAGTCTGATGATGATGATGATGATGATGATG 31201
QY 3691 CCTGCTTCTCAGCTCTTATATCTTAAGACACAGCATATCTCTTGAAGAAATTAACAGC 3750

DB 31202 CCTCGTTCTCAGCTCTTATATTCCTTAAGACACAGCATCATCTCTAGTAATACAC 31261
QY 3751 TAATTGGCAGTGAAGCCGAGATCGACCACTGCACCCCTGCTGGGAGACAGTGAAGCT 3810
DB 31262 TAATTGGCAGTGAAGCCGAGATCGACCACTGCACCCCTGCTGGGAGACAGTGAAGCT 31321
QY 3811 TTGTCTCTATTTACAAAAGAAAAGAAAAGAAAAGTAACCTAACCTCA 3857
DB 31322 TTGTCTCTATTTACAAAAGAAAAGAAAAGAAAAGTAACCTAACCTCA 31368

RESULT 7
AAS63229

ID AAS63229 standard; cDNA; 1101 BP.

AC AAS63229;

DT 29-JAN-2002 (first entry)

XX DNA encoding human XPF (ERCC4) active peptide fragment.

XX DNA mutation-binding protein; nuclease: DNA mismatch; cancer; PCR primer;
KW DNA damage; human xeroderma pigmentosum complementation group; XPF; XPA;
KW XPC; XPE; ERCC4; human Muts homolog 2; hMSH2; Muts; Nuc; Muty; Fpg; ss;
KW Fapy-DNA glycosylase; uracil DNA glycosylase; ung; TDG; xhda gene; Uvr A;
KW A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;
KW thymine DNA-glycosylase; Uvr B; Uvr C; nch gene; nfo gene; exonuclease;
endonuclease.

XX Homo sapiens.

XX W0200173079-A2.

XX 04-OCT-2001.

XX 26-MAR-2001; 2001WO-US09700.

XX 28-MAR-2000; 2000US-192764P.

XX 29-AUG-2000; 2000US-0650855.

XX (REGC) UNIV CALIFORNIA.

XX Mc Cutchen-maloney SL;

XX WPI: 2001-656920/75.

XX P-PSDB; AAU69742.

XX Recombinant chimeric protein, useful for detecting and quantifying DNA
PT mutations, e.g. in disease diagnosis, comprises mutation-binding
PT protein and nuclease -

XX PS Disclosure; Page 65; 128pp; English.

XX Sequences AAS63224-AAS63305 represent cDNA encoding and PCR primers used
CC to amplify cDNA encoding proteins which can be used in the synthesis of
CC chimeric proteins comprising a DNA mutation-binding protein, a linker and
CC a nuclease, by recombinant technology. The chimeric proteins are useful
CC for detection, quantification and mapping of DNA sequence variations
CC including mutations, for example, caused by damage and mismatches. The
CC proteins are able to bind to the site of the DNA mutation and cut it out
CC of the molecule. This is useful for early diagnosis of cancer and other
CC diseases. The proteins used in the invention include human XPF (or
CC ERCC4), human xeroderma pigmentosum complementation groups A, C and E
CC (XPA, XPC and XPE), human Muts homolog 2 (hMSH2), Serratia marcescens
CC nuclease (Nuc), Thermus thermophilus Muts, Escherichia coli Fapy-DNA
CC glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific adenine
CC glycosylase (Muty), synthetic T4 endonuclease V (T4 endo V), thymine
CC DNA-glycosylase (TDG), E. coli Uvr A, B and C, and E. coli endonucleases
CC and exonucleases.

XX Sequence 1101 BP; 324 A; 230 C; 246 G; 301 T; 0 other;

Query Match

28.5%; Score 1101; DB 22; length 1101.

Best Local Similarity 100.0%; Pred. No. 1,5e-271;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 ATGGCCCCGCTGCTGGAGTAGAGAGCAGACAGCTGGTGTGAACTGCTCGACACTGACGG 105
DB 1 ATGGCCCCGCTGCTGGAGTAGAGAGCAGACAGCTGGTGTGAACTGCTCGACACTGACGG 60
QY 106 CTAGTGTGTGCCCCCGGGGGCTCGGGGGGACCGGGCTCTACCACTTCTCCAGCTG 165
DB 61 CTAGTGTGTGCCCCCGGGGGCTCGGGGGGACCGGGCTCTACCACTTCTCCAGCTG 120
QY 166 CACTGCCACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225
DB 121 CACTGCCACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 226 TTTATCATCAGCTGAAGATAGAAGAGTTGAACACCTCCCTCCGCTGTAAACAAATGA 285
DB 181 TTTATCATCAGCTGAAGATAGAAGAGTTGAACACCTCCCTCCGCTGTAAACAAATGA 240
QY 286 ATCACAAGCAACAGTGGCTATGAAGTTTACACCAAGGTGGTATATTGGGACAACT 345
DB 241 ATCACAAGCAACAGTGGCTATGAAGTTTACACCAAGGTGGTATATTGGGACAAAGT 300
QY 346 AGGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 405
DB 301 AGGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 406 TTGCTGTATAGAGCCCAACAGATATGAGCTTGTGCAAGACATTCATCTGGCCCTC 465
DB 361 TTGCTGTATAGAGCCCAACAGATATGAGCTTGTGCAAGACATTCATCTGGCCCTC 420
QY 466 TTTCGGCAGAAAACAAGCTGTTTATTAAGCTTTCACAGCAAAVGGTGGCTTT 525
DB 421 TTTCGGCAGAAAACAAGCTGTTTATTAAGCTTTCACAGCAAAVGGTGGCTTT 480
QY 526 GATACCTGTTTTTGTGATGGAAGAGTGAAGAACTTTTTGTGAGAAACTGTAT 585
DB 481 GATACCTGTTTTTGTGATGGAAGAGTGAAGAACTTTTTGTGAGAAACTGTAT 540
QY 586 CTGTGGCCAAAGTTCATGATGAGCAATTAATCTATTTTGAACAGCAAAAGCTGAAGT 645
DB 541 CTGTGGCCAAAGTTCATGATGAGCAATTAATCTATTTTGAACAGCAAAAGCTGAAGT 600
QY 646 GTAGAAATCCATGTTTCTATGACACCTACCATCTTGTCTATACAGACTGCTATACAG 705
DB 601 GTAGAAATCCATGTTTCTATGACACCTACCATCTTGTCTATACAGACTGCTATACAG 660
QY 706 ATTTTAAATGATGCTTAAGAACTAAATGCTATTAACCATACCTGCTGAAGTGAAGAT 765
DB 661 ATTTTAAATGATGCTTAAGAACTAAATGCTATTAACCATACCTGCTGAAGTGAAGAT 720
QY 766 TTTATCTTTAAGAAATGCTATTAAGAACTTTGACAAAGCAATCCGCCATTTCTGGAT 825
DB 721 TTTATCTTTAAGAAATGCTATTAAGAACTTTGACAAAGCAATCCGCCATTTCTGGAT 780
QY 826 CTTTGTGGCAGCAGCTTGGAGCCAAAGCTAAATCTTTATGAGATTTGAAGATTTA 885
DB 781 CTTTGTGGCAGCAGCTTGGAGCCAAAGCTAAATCTTTATGAGATTTGAAGATTTA 840
QY 886 CGAATCTTGTGCTGATGCTCTCTCAGTATGATGCTGCAATTTCTTAATCTTCTGGA 945
DB 841 CGAATCTTGTGCTGATGCTCTCTCAGTATGATGCTGCAATTTCTTAATCTTCTGGA 900
QY 946 TCTCTGAGACCAACGAAAAAGCTTTTGGTCACAAATTCAGTGGCGTTCCTTGACTCC 1005
DB 901 TCTCTGAGACCAACGAAAAAGCTTTTGGTCACAAATTCAGTGGCGTTCCTTGACTCC 960
QY 1006 AGCAGCTCGATGTTTAAATGCTCGAGCAAGGTTTATCATCTTCAGATGCAAAATG 1065
DB 961 AGCAGCTCGATGTTTAAATGCTCGAGCAAGGTTTATCATCTTCAGATGCAAAATG 1020
QY 1066 AGTAAAAAGAAAAAATATCTGAAAAAATGAAATTAAGAGGCAAGAAAAAAG 1125
DB 1125 AGTAAAAAGAAAAAATATCTGAAAAAATGAAATTAAGAGGCAAGAAAAAAG 1080

DB 1021 AGTAAAAAGAAAAATTCGAAAAATTAAGAGGGGAACAAAAAG 1080

QY 1126 GAACGTGCTCAGAAAGCAAC 1146
|||||

DB 1081 GAACGTGCTCAGAAAGCAAC 1101

RESULT 8

ABL87494/c
ID ABL87494 standard; cDNA; 601 BP.
XX
AC ABL87494;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:10472.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN M0200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI: 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
XX
PS Claim 1; SEQ ID 10472; 4899p; English.

CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (SI) from the 10912 nucleotide sequences as given in ABL77003 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (SI) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
SQ Sequence 601 BP; 180 A; 134 C; 118 G; 166 T; 3 other.

Query Match 6.6%; Score 256.2; DB 24; Length 601;
Best Local Similarity 97.7%; Pred. No. 2.8e-55;
Matches 258; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1561 GAATTAAGCAGTACCCAGAAAGTCCCGGAGAAATTAGCAATTTGATGATGA 1620
|||||

DB 601 GAATTAAGCAGTACCCAGAAAGTCCCGGAGAAATTAGCAATTTGATGATGA 542

QY 1621 AATTGTCATGCGATGCTGCTTTGGAATCCGAAAGAACCCCTCATTCATCCGCG 1680
|||||

DB 541 AATNTGTCATGCGATGCTGCTTTGGAATCCGAAAGAACCCCTCATTCATCCGCG 482
|||||

QY 1681 CTCTGGGTTGACGAGCCACCTATGCTGTGACAAAGGTTACTACATGAAGTGAGCCAGA 1740
|||||

DB 481 CTCTGGGTTGACGAGCCACCTATGCTGTGACAAAGGTTACTACATGAAGTGAGCCAGA 422
|||||

QY 1741 TACGTGTTCTTTATGACGAGAGCTAACCTTTGTCGCGACGCTTGAATTTACAGGCG 1800
|||||

DB 421 TACGTGTTCTTTATGACGAGAGCTAACCTTTGTCGCGACGCTTGAATTTACAGGCG 362
|||||

QY 1801 AGTAGCCTGGGAAACCTCTGAGG 1824
|||||

DB 361 AGTAGCCTGGGAAACCTCTGAGG 338
|||||

RESULT 9

ABL06559
ID ABL06559 standard; cDNA; 2995 BP.
XX
AC ABL06559;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 14159.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN M0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR P-PSDB: ABB62456.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 14159; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 2995 BP; 775 A; 749 C; 807 G; 664 T; 0 other.

Query Match 6.6%; Score 253.2; DB 23; Length 2995;
Best Local Similarity 46.9%; Pred. No. 3.7e-54;
Matches 1270; Conservative 0; Mismatches 1303; Indels 135; Gaps 10;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:45:25 ; Search time 3612.58 Seconds

(without alignments)
17291.221 Million cell updates/sec

Title: US-09-919-497-14

Perfect score: 3857

Sequence: 1 ggaagagctccatgagtc.....agaatacaactagctca 3857

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1062.2	27.5	1083	11	BC020741 Homo sapi
2	780.6	20.2	818	13	B1522552 603175243
3	671.8	17.4	748	12	BG620282 602618570
4	653.8	17.0	741	12	BG181154 602329292
5	610.4	15.8	1119	11	AK018375 Mus muscu
6	537.8	13.9	544	14	BQ013114 UI-1-BC1P

c	7	537.4	13.9	581	10	AW977575	AW977575 EST898684
c	8	530.6	13.8	683	10	AM242081	AM242081 x021c12.x
c	9	520.6	13.5	530	14	BQ011470	BQ011470 UI-1-BC1P
c	10	513.6	13.3	829	13	B1555512	B1555512 603239318
c	11	499.6	13.0	515	14	BQ310815	BQ310815 OV3-BN004
c	12	499	12.9	524	13	BM671280	BM671280 UI-E-CQ1-
c	13	485.6	12.6	923	14	BE818393	BE818393 RC5-BN028
c	14	482.2	12.5	923	14	BQ944411	BQ944411 AGENCOURT
c	15	467	12.1	596	10	AV685090	AV685090 AV685090
c	16	466	12.1	466	14	BM710111	BM710111 UI-E-CQ1-
c	17	457.4	11.9	459	9	AL705565	AL705565 DKFZP6861
c	18	456.2	11.8	509	12	BE837466	BE837466 RC2-FN009
c	19	453.6	11.8	460	9	AT025538	AT025538 ov76d11.x
c	20	452.2	11.7	752	12	BG724387	BG724387 602692867
c	21	452	11.7	464	9	AA256859	AA256859 zr83a11.r
c	22	451.4	11.7	651	10	BB640850	BB640850 BB640850
c	23	447.4	11.6	708	9	AJ447405	AJ447405 AJ447405
c	24	437	11.3	449	9	AA291199	AA291199 z547c05.r
c	25	436.2	11.3	463	9	AT018046	AT018046 ov64c11.s
c	26	433.6	11.2	641	9	AJ447403	AJ447403 AJ447403
c	27	422.2	10.9	887	10	BE565954	BE565954 601328232
c	28	421.4	10.9	434	9	AA770518	AA770518 ah74d10.s
c	29	416.6	10.8	640	12	BG866533	BG866533 602785491
c	30	407.8	10.6	418	9	AT431784	AT431784 t125b12.x
c	31	400.8	10.4	715	13	B1459712	B1459712 603200751
c	32	395.6	10.3	425	9	AA774566	AA774566 a127c07.s
c	33	385.6	10.0	403	10	AV692790	AV692790 AV692790
c	34	383.2	9.9	801	9	AJ447404	AJ447404 AJ447404
c	35	380.4	9.9	654	9	AJ447402	AJ447402 AJ447402
c	36	358.8	9.3	374	12	BE818447	BE818447 RC5-BN028
c	37	355	9.2	523	10	AM271424	AM271424 x514d04.x
c	38	340.8	8.8	471	10	BB856543	BB856543 BB856543
c	39	339.6	8.8	357	9	AA638976	AA638976 ns01b08.s
c	40	322	8.3	583	13	B1693586	B1693586 603341969
c	41	312.8	8.1	668	10	AW921189	AW921189 EST352493
c	42	309.4	8.0	501	9	AA436963	AA436963 zv72f08.r
c	43	296	7.7	296	9	AA255461	AA255461 zr83a11.s
c	44	295.2	7.7	312	9	AA723776	AA723776 ab62f10.s
c	45	282	7.3	470	9	AT013171	AT013171 EST207846

ALIGNMENTS

RESULT 1
LOCUS BC020741 1083 bp mRNA linear HTC 22-JAN-2002
DEFINITION Homo sapiens, Similar to excision repair cross-complementing rodent repair deficiency, complementation group 4, clone IMAGE:4733192, mRNA.

ACCESSION BC020741
VERSION BC020741.1 GI:18089125

KEYWORDS HTC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 1083)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbbs-remail.nih.gov

Tissue Procurement: CLOMTECH

cDNA Library Preparation: CLOMTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMN)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mdedpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers
R. M.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINk at: <http://image.lnh.gov>
Series: IRAL Plate: 37 Row: h Column: 11
This clone has the following problem: frame shifted.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4732192"
/risue_type="Placenta"
/clone_id="NH.MGC.79"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
BASE COUNT      322 a      229 g      235 g      297 t
ORIGIN

```

27.5%; Score 1062.2; DB 11; Length 1083;

Best Local Similarity	98.8%;	Pred No. 9,36-214;
Matches 1070; Conservative	0;	Mismatches 13; Indels 0; Gaps 0;

QY	33	TCGAGGATTTC	CCATGGCGCCGCTGCTG	AGATACGAGCACACCTG	GTGCTGGAACTGCT	92		
Db	1	TCGAGGGATTTC	CCATGGCGCGCTGCTG	AGATAGAGCCGACACTG	GTGCTGGAACTGCT	60		
OY	93	CGACACTAGCGGGCT	AGTAGTGTGCGGCCGCGGGCTG	CGGGCGGACCGGGCTGCT	CTATACCA	152		
Db	61	CGACACTAGCGGGCT	AGTAGTGTGCGGCCGCGGGCTG	CGGGCGGACCGGGCTGCT	CTATACCA	120		
OY	153	CTTTCTCCAGCTG	CACACGCCAGCCTGCTG	TGCTGCTCAACACGACGCGGC		212		
Db	121	CTTTCTCCAGCTG	CACACGCCAGCCTGCTG	TGCTGCTCAACACGACGCGGC		180		
OY	213	CGAGGAGGAGTAT	TTATCAATCAGCTG	AGATAGAAAGAGTTG	AGACACTCCCTGCGCG	272		
Db	181	CGAGGAGGAGTAT	TTATCAATCAGCTG	AGATAGAAAGAGTTG	AGACACTCCCTGCGCG	240		
OY	273	TGTAACTAAATG	ATCACAACGCAACGCGGT	ATGAAAGTTTACACACAGAGTG	TAT	332		
Db	241	TGTAACTAAATG	ATCACAACGCAACGCGGT	ATGAAAGTTTACACACAGAGTG	TAT	300		
OY	333	ATTTCGCAAGTAG	AGATACTTGTGTTG	ATCTTTGACTGAT	AGAAATACCTTCAGATT	392		
Db	301	ATTTCGCAAGTAG	AGATACTTGTGTTG	ATCTTTGACTGAT	AGAAATACCTTCAGATT	360		
OY	393	AATTACTGGCAT	CTTTGGTGTATAG	AGCCACACGATAT	ATGAGCTTTGTCAAGACATT	452		
Db	361	AATTACTGGCAT	CTTTGGTGTGTATAG	AGCCACACGATAT	ATGAGAGCTTTGTCAAGACATT	420		
OY	453	CATCTGGGCGCT	TTGGCCGAGAAAACAA	CGAGTTTATTAACCTTC	CACAGCAA	512		
Db	421	CATCTGGGCGCT	TTGGCCGAGAAAACAA	CGAGTTTATTAACCTTC	CACAGCAA	480		
OY	513	TGCTGTTCCTTGA	TACTGTTTGTGATG	GTGAAGAGTATG	GAATCTTTTGT	572		
Db	481	TGCTGTTCCTTGA	TACTGTTTGTGATG	GTGAAGAGTATG	GAATCTTTTGT	540		
OY	573	GAGGAAGCTG	TATCTGTGGCCAA	GGTTCATGACGAT	AACTCATTTT	TAGACAGCA	632	
Db	541	GAGGAAGCTG	TATCTGTGGCCAA	GGTTCATGACGAT	AACTCATTTT	TAGACAGCA	600	
OY	633	CAAACTTAAGTTG	ATAGAAATCCANTG	TTATATACACCTA	CCATGCTGTAT	ACAGAC	692	
Db	601	CAAACTTAAGTTG	ATAGAAATCCANTG	TTATATACACCTA	CCATGCTGTAT	ACAGAC	660	
OY	693	TGCTTACTG	CATTTTAAATG	CTATGCTAAAGAACT	AAATAGCATAA	CCATGCT	752	
Db	661	TGCTTACTG	CATTTTAAATG	CTATGCTAAAGAACT	AAATAGCATAA	CCATGCT	720	
OY	753	TGAAGTGA	AGATTTATCTTT	AGAAATGCTATT	GTGAACCTTT	TGACAAAGACAT	CGC	812

Db	721	TGAAGTGGAGAGATTATCTTAGAAGAAATGCTATTGGAAACCTTTTGACAAGACAAATCCG	780
OY	813	CCATTATCTGCATCTTTGTGTGGCACCACCTTGGACCAGACTAATCTTAAGTGCAGGA	872
Db	781	CCATTATCTGCATCTTTGTGTGGCACCACCTTGGACCAGACTAATCTTAAGTGCAGGA	840
OY	873	TTTGAAGATATATACGAACCTTGTGCTCAGTATCTCTCAGTATGATTTGTGCACATTTCT	932
Db	841	TTTGAAGATATATACGAACCTTGTGCTCAGTATCTCTCAGTATGATTTGTGCACATTTCT	900
OY	933	TAACTTCTGCAATCTCTGAGAGACAAGCAAGAAAACCTTTTGTGCAGAAATTCAGGTTGGCT	992
Db	901	TAACTTCTGCAATCTCTGAGAGACAAGCAAGAAAACCTTTTGTGCAGAAATTCAGGTTGGCT	960
OY	993	GTTTCTTGACCTCAGCACCCTCGATCTTTATTAATGCTCGAGCAACGCTTATCATCTTCC	1052
Db	961	GTTTCTTGACCTCAGCACCCTCGATCTTTATTAATGCTCGAGCAACGCTTATCATCTTCC	1020
OY	1053	AGATGCCAAAATGAGTAAAAAAGAAAAAATATCTGAAAAAATGCAAAATTTAAAGAAGGGGA	1112
Db	1021	AGATGCCAAAATGAGTAAAAAAGAAAAAATATCCAAAAAATTTAAAGAAGGGGA	1080
OY	1113	AGA 1115	
Db	1081	AAA 1083	

[illegible]

KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homologous Protein	Sequence Identity	Accession Number
Homo sapiens	Chordin	99	U01452
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi	Chordin	99	U01452

REFERENCE
1 (bases 1 to 818)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Life Technologies, Inc.
Email: cgabbs-remail.nin.gov

CDNA library prepared by: The I.M.A.G.E. Consortium (LLNL)
CDNA Sequencing by: Invitae Genomics, Inc.

UMA sequencing by: mycote genomics; InciClone distribution: MGC clone distribution information can be found through the T M A G E Consortium/INi. at:

found through the Linn. Soc. London
http://image.llnl.gov
plate: IAM11604 row: e column: 11

High quality sequence stop: 802.
Location/Qualifiers

SOURCE

```
/db_xref="taxon:9606"  
/clone="IMAGE:5239666"
```

```
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
```

```
/note="Organ: brain; Vector: pcMV-SpORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of
```

fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range

0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics cracking code 017. Note this is a NIH_MGC library."

BASE COUNT	241 a	158 c	167 g	252 t
------------	-------	-------	-------	-------

QY 513 TCGTGTGCTTTGATACCTGTTTGTCTCATGTGGAAGAGATGATGAATCTTTTGT 572
 Db 482 TCGTGTGCTTTGATACCTGTTTGTCTCATGTGGAAGAGATGATGAATCTTTTGT 540
 QY 573 GAGGAACCTATCTGTGGCCAAAGTTCCATGTAGCAGTAACTCATTTTGTAGACGCA 632
 Db 541 GAGGAACCTATCTGTGGCCAAAGTTCCATGTAGCAGTAACTCATTTTGTAGACGCA 600
 QY 633 CAACCTGAAGTGTAGAAATTCATGTTTGTATGACACCTGATGCTTGTATGACGAC 692
 Db 601 CAACCTGAAGTGTAGAAATTCATGTTTGTATGACACCTGATGCTTGTATGACGAC 660
 QY 693 TCGTATACCTGACATTTTAAATGATGCTTAAAGAACTAAATGCCATACCATGCT 752
 Db 661 TGTATCTGTGACATTTTAAATGATGCTTAAAGAACTAAATGCCATACCATGCT-- 718
 QY 753 TGAAGTGGAGATTTATCTTTAGAAATG 781
 Db 719 TTGAGGTGCACGATTTATCTTTAGAAAGG 747

RESULT 4
 LOCUS BG181154 741 bp mRNA linear EST 06-FEB-2001
 DEFINITION 602329292F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4430639 5',
 mRNA sequence.
 ACCESSION BG181154
 VERSION BG181154.1 GI:12687857
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: DCTP/DRP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Note: Organ: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM10184 row: g column: 24
 High quality sequence stop: 677.

FEATURES
 source
 Location/Qualifiers
 1..741

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4430639"
 /clone_lib="NIH_MGC_91"
 /tissue_type="adenoecarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Note: this is a NIH_MGC Library."
 BASE COUNT 218 a 151 c 150 g 222 t
 ORIGIN

Query Match 17.0%; Score 653.8; DB 12; Length 741;
 Best Local Similarity 96.9%; Pred.No.1.5e-127;
 Matches 720; Conservative 0; Mismatches 17; Indels 6; Gaps 5;
 QY 248 AAGAGCTTGAACACCTCCCTCGCGTGTAAACAATGAATGACAGACAGTGGCTAG 307
 Db 1 AAGAGCTTGAACACCTCCCTCGCGTGTAAACAATGAATGACAGACAGTGGCTAG 60

QY 308 AAGTTTACACAAAGTGTGTTATATTTCGACAGTAGATACTGTGGTGAATCTCT 367
 Db 61 AAGTTTACACAAAGTGTGTTATATTTCGACAGTAGATACTGTGGTGAATCTCT 120
 QY 368 TGAATGATAGAAATCTTCAAGATTTAAATCTGAGCATCTTGGTGTATAGACCCACGAA 427
 Db 121 TGAATGATAGAAATCTTCAAGATTTAAATCTGAGCATCTTGGTGTATAGACCCACGAA 180
 QY 428 TAAATGAGCTTGTCAAGAAAGCAATCACTTGGGCTCTTCCGCNCAAAAACAAACGCTG 487
 Db 181 TAAATGAGCTTGTCAAGAAAGCAATCACTTGGGCTCTTCCGCNCAAAAACAAACGCTG 240
 QY 488 GTTTTATTAAGCTTTCACAGACAAATGCTTGTGCTTGTATGATGCTTTTGTGATGCTG 547
 Db 241 GTTTTATTAAGCTTTCACAGACAAATGCTTGTGCTTGTATGATGCTTTTGTGATGCTG 300
 QY 548 AAGAGCTGATGAGAAATCTTTTGTGAGGAAGTATCTGTGGCCAA-GCTTCATGTA 606
 Db 301 AAGAGCTGATGAGAAATC-TTTTGTGAGGAAGTATCTGTGGCCAAAGGTTCCATGTA 359
 QY 607 GCAGTAACTCATTTTGTAGAAAGCAACCAACCTGAAGTTGTAGAAATTCATGTTTGTATG 666
 Db 360 GCAGTAACTCATTTTGTAGAAAGCAACCAACCTGAAGTTGTAGAAATTCATGTTTGTATG 419
 QY 667 ACACCTACCACTGCTTGTATACAGACAGCTGATCTGACATTTTAAATGCAATGCTTAAAG 726
 Db 420 ACACCTACCACTGCTTGTATACAGACAGCTGATCTGACATTTTAAATGCAATGCTTAAAG 479
 QY 727 GAACCTAAATGCCATTAACCCATGCTTGAAGTGAAGATTTATCTTTAGAAATGCTATT 786
 Db 480 GAACCTAAATGCCATTAACCCATGCTTGAAGTGAAGATTTATCTTTAGAAATGCTATT 539
 QY 787 GGAACCTTTTGTAGAAAGCAATCCGCCATTAATCTGATGCTTGTGCGCCAGCCTTGA 846
 Db 540 GGAACCTTTTGTAGAAAGCAATCCGCCATTAATCTGATGCTTGTGCGCCAGCCTTGA 598
 QY 847 GCCAAGACTAAATCTTGTAGTTCAGGA-TTGAAGATATTGCAACTTGTGTCAGATATCT 905
 Db 599 GCCAAGACTAAATCTTGTAGTTCAGGA-TTGAAGATATTGCAACTTGTGTCAGATATCT 658
 QY 906 CTCTCAGTATGATTTGTCTACAT--TTCTTAATCTTGTGAATCTGTGAGACGACGAA 963
 Db 659 CTCTCAGTATGATTTGTCTACAT--TTCTTAATCTTGTGAATCTGTGAGACGACGAA 718
 QY 964 AAGCTTTTGTGTCAGAAATTCAGG 986
 Db 719 CAACGCTTTTGGCAGAAATCCCG 741

RESULT 5
 LOCUS AK018375 1119 bp mRNA linear HNC 19-JAN-2002
 DEFINITION Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched
 library, clone:8430401P11:excision repair cross-complementing
 rodent repair deficiency, complementation group 4, full insert
 sequence.

ACCESSION AK018375
 VERSION AK018375.1 GI:12858047
 KEYWORDS HNC, CAP trapper.
 SOURCE Mus musculus (Strain: C57BL/6J) 16 days embryo lung cDNA to mRNA,
 clone_lib:RIKEN full-length enriched mouse cDNA library
 clone:8430401P11.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 AUTHORS Carninci, P., and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

1 Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

2 20499374
11042159

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,Y., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,S., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasliwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matabiki,K.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,T., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multipipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

4 20530913
11076861

5 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Aikawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kusukawa,T., Saito,R.,
Kadoi,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,M., Gaasterland,T., Gissi,C., Kling,B., Kochwa,T.,
Kuehl,P., Lewis,S., Matsuo,T., Nikaido,T., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarrelli,R., Barsh,G., Blake,T., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustinich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamaya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazarrelli,J., Mombaerts,P., Noriome,P.,
R.ing,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-Oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kotsuki,S.,
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

6 21085660
11217851

7 5 (bases 1 to 1119)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Aikawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hangaki,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,O., Kojima,Y., Kono,H., Kouda,M., Koya,S.,
Kunihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numata,S., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer

[illegible]

```

Db 368 GATTCTTGACAGTAGAATACCTTCAGATTATATACTGGCATCCTGCTGTACAGGCA 427
Qy 421 CACAGATATATCGATCTTGTCTCAAGAACATTCATCTTGCGCTCTTGGCCAGAAAAC 480
Db 428 CACAGATATATGAGTCCTGCGCAGGAAGCCTTCATCTGCGCTCTTGGCCAGAAAAC 487
Qy 481 AAACGAGTCTTTTAAAGCTTTCACAGCAATGCTGTCCTTGTATCTGCTTTTGT 540
Db 488 AACGGCGGCTTCATCAAGCTTTCACCGACACAGCTGTCCTTGTACACTGCTTTGT 547
Qy 541 CATGTGGAAGAGTGTAGAGAAATCTTTTGTGAGGAACCTGTATCTGTGGCCAGGTC 600
Db 548 CACGTGGAAGAGTGTAGAGAAATCTTTTGTGAGGAACCTGTATCTGTGGCCAGGTC 607
Qy 601 CACTGACAGTAACTCATTTTATAGAACACCAACCTGAATCTGTAGAAATCCATGTT 660
Db 608 CACTGACAGTAACTCATTTTGTGAGAACACCAACCTGAATCTGTAGAGATTCACGTC 667
Qy 661 TCTATGACACCTCATCTGCTTCTATACAGACGTATCTGTGACATTTTAAATGATGT 720
Db 668 TCCATGACCTCCGCCATGCTGTGCATTCACAGGCCATCTACTAGACATCTTAATGATGT 727
Qy 721 CTAAAGAACTAAATGCAATACCCATGCTTGAAGTGAAGATTTATCTTTAGAAAT 780
Db 728 CTGAAGAACTGAATGTGACAAACCCGCTCACTGAAGTAGAATTTGTACATAAAG 787
Qy 781 GCTATGGAAGAACCTTTTGACAG 804
Db 788 GCTCTTGGGAAGCAATTCGACAG 811

```

```

RESULT 5
LOCUS B0013114/c 544 bp mRNA linear EST 26-MAR-2002
DEFINITION UI-1-BCIP-ay1-f-03-0-UI s1 NCI-CGAP_P13 Homo sapiens cDNA clone
ACCESSION B0013114
VERSION B0013114.1 GI:19738015
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 544)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Steven Brown
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
CDNA Distribution: Clome distribution information can be obtained
from Dr. M. Bento Soares, Bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLY-A=yes.

```

```

FEATURES
source
1..544
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-1-BCIP-ay1-f-03-0-UI"
/clone_lib="NCI-CGAP_P13"
/tissue_type="Placenta"
/dev_stage="8-9 weeks"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Placenta; Vector: pTV73-Pac (Pharmacia) with
a modified polylinker; Site: 1; Ecor I; Site 2; Not I;
NCI-CGAP_P13 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double

```

stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTV73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tags for this library are GA, AGGA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu

TMG_Lib-UI-1-BCIP
TAG_TISSUE=Placenta human 8 week
TAG_SEQ=GA*

```

BASE COUNT 163 a 106 c 101 g 174 t
ORIGIN
Query Match 13.9%; Score 537.8; DB 14; Length 544;
Best Local Similarity 99.6%; Pred. No. 4,7e-103;
Matches 539; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
543 TGTGAAAGAGTGTAGAAATCTTTTGTGAGAAACTGTATCTGCGCAAGTTCCA 602
544 TGTGAAAGAGTGTAGAAATCTTTTGTGAGAAACTGTATCTGCGCAAGTTCCA 485
603 TGTAGCACTAACTCATTTTATAGACAGCAAACTGAAGTTGAGAAATCCATGTTTC 662
484 TGTAGCACTAACTCATTTTATAGACAGCAAACTGAAGTTGAGAAATCCATGTTTC 425
663 TATGACACCTACCATGCTGCTATACAGACTGCTATACAGCAATTTAAATGATGTC 722
424 TATGACACCTACCATGCTGCTATACAGACTGCTATACAGCAATTTAAATGATGTC 365
723 AAAGAACTAAATGCAATACCCATGCTTGAAGTGAAGATTTATCTTTAGAAAATGC 782
364 AAAGAACTAAATGCAATACCCATGCTTGAAGTGAAGATTTATCTTTAGAAAATGC 305
783 TATTGGAAGAACTTTTGACAGCAAACTCCGCTATATCTGAGATCTTTGTGGACACGCT 842
304 TATTGGAAGAACTTTTGACAGCAAACTCCGCTATATCTGAGATCTTTGTGGACACGCT 245
843 TGGAGCCAAAGACTAAATCCCTAGTTCAGATTTGAAGATTTAGCAACTTTGCTGAGTA 902
244 TGGAGCCAAAGACTAAATCCCTAGTTCAGATTTGAAGATTTAGCAACTTTGCTGAGTA 185
903 TCTCTCTCACTATGATTTGTCTCAATTTTAAATCTTTGGAATCTTGTGAGACACGCA 962
184 TCTCTCTCACTATGATTTGTCTCAATTTTAAATCTTTGGAATCTTGTGAGACACGCA 125
963 AAAAGCTTTTGTGACAGATTCAGTTGGCTGCTTCTTGAATCCAGACCTCGATGTTAT 1022
124 AAAAGCTTTTGTGACAGATTCAGTTGGCTGCTTCTTGAATCCAGACCTCGATGTTAT 65
Qy 1023 AAATGCTCGAAGAGGTTTATCATCTTCCAGATGCCAAATGATTAAGAAAAAT 1082
Db 64 AAATGCTCGAAGAGGTTTATCATCTTCCAGATGCCAAATGATTAAGAAAAAT 5
Qy 1083 A 1083
Db 4 A 4

```

```

RESULT 7
LOCUS AW977575/c 581 bp mRNA linear EST 02-JUN-2000
DEFINITION EST389684 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW977575
VERSION AW977575.1 GI:8168828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and

```


Db 266 GCTCTTGGTGGTCCAGGTCACAGTACTGATTTGTGCAAGTGAATGACCAGCAATGTTCCAG 207

Qy 1270 CTGAGACAGATATATACACTCTTGTGGACGGAGCCTTCTTATTTAGGCTCTACAGAAACC 1329

Db 206 CTGACAGACTATATACACTCTTGTGGACGGAGCCTTCTTATTTAGGCTCTACAGAAACC 147

Qy 1330 TTGTGAAAGGA--TGGCAAGAGCTGGAAGAGTGTGA---TGAAATTTTGGAAAGAGAC 1383

Db 146 TTGTGAAAGGATAGCAAGAGCTGGAAGAGTGTGAAGTGAAGTGAAGAGAGAC 87

Qy 1384 ACTTCAAA-GAGAAATTAG-GAAATCTCACAAAGACCTTAAGACCCCAAAACAAAGAAC 1441

Db 86 ACTTCAAAAGAAATTAGGAAATCTCACAAAGACCTTAAGACCCCAAAACAAAGAAC 27

Qy 1442 GGGCTTCTACCAAGAAAGAACCTTC 1467

Db 26 GGGCTTCTACCAAGAAAGAACCTTC 1

RESULT 9

LOCUS B0011470 530 bp mRNA linear EST 26-MAR-2002

DEFINITION UI-1-BC1P-ase-e-04-0-UI.s1 NCI_CGAP_P13 Homo sapiens cDNA clone

ACCESSION B0011470

VERSION UI-1-BC1P-ase-e-04-0-UI 3', mRNA sequence.

KEYWORDS B0011470.1 GI:19736371

SOURCE EST.

ORGANISM human.

Human sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 530)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Steven Brown

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLY-A=yes.

FEATURES

source

1..530

Location/Qualifiers

1..530

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="UI-1-BC1P-ase-e-04-0-UI"

/clone_1lb="NCI_CGAP_P13"

/tissue_type="Placenta"

/dev_stage="8-9 weeks"

/lab_host="PH10B (Life Technologies)"

/note="Organ: Placenta; Vector: pTR73-Pac (Pharmacia) with a modified polylinker. Site_1: EcoR I; Site_2: Not I; NCI_CGAP_P13 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTR73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGCA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu

TAG_LIB=UI-1-BC1P

TAG_TISSUE=Placenta human 8 week

TAG_SEQ=GA"

BASE COUNT 160 a 99 c 101 g 170 t

ORIGIN

Query Match 13.5%; Score 520.6; DB 14; Length 530;

Best Local Similarity 99.2%; Pred. No. 2e-99;

Matches 523; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 557 TGAGAAATTTTTTTTGAGAAATGTATCTGTGCGCAGGTTCCATGTAGCAATTAAC 616

Db 530 TGAGAAATTTTTTTTGAGAAATGTATCTGTGCGCAGGTTCCATGTAGCAATTAAC 471

Qy 617 CATTTTACAGACAGCAAAACCTGAAGTTGTGAATTCATTTTCTATGACACTACCA 676

Db 470 CATTTTACAGACAGCAAAACCTGAAGTTGTGAATTCATTTTCTATGACACTACCA 411

Qy 677 TCGTGTCTATACAGACTGTATACAGATTTTAATGATGTCTTAAGAACTAAAT 736

Db 410 TCGTGTCTATACAGACTGTATACAGATTTTAATGATGTCTTAAGAACTAAAT 351

Qy 737 GCCATTAACCCATCGCTTGAAGTGAAGATTTTACCTTTAGAAATGCTATTGAAAACTT 796

Db 350 GCCATTAACCCATCGCTTGAAGTGAAGATTTTACCTTTAGAAATGCTATTGAAAACTT 291

Qy 797 TTGACAAAGCAATCCGCCATTTATCTGATCTTTTGTGCGCAGCTTGGACCAAGACTA 856

Db 290 TTGACAAAGCAATCCGCCATTTATCTGATCTTTTGTGCGCAGCTTGGACCAAGACTA 231

Qy 857 AATCCTTAGTTCAGATTTGAAGATTTTGAAGATTTTGTGCGCAGTATCTCTCACTATG 916

Db 230 AATCCTTAGTTCAGATTTGAAGATTTTGAAGATTTTGTGCGCAGTATCTCTCACTATG 171

Qy 917 ATGTGTCAATTTCTTAATCTCTGGAATCTGTGAGCAACGCAAAAGCTTTTGGTC 976

Db 170 ATGTGTCAATTTCTTAATCTCTGGAATCTGTGAGCAACGCAAAAGCTTTTGGTC 111

Qy 977 AGAATTCAGTGGCTGCTTTCTGACTCCAGACCTCGATGTTTAATGCTCGAGCA 1036

Db 110 AGAATTCAGTGGCTGCTTTCTGACTCCAGACCTCGATGTTTAATGCTCGAGCA 51

Qy 1037 GGGTTATCATCTCCAGATGCCAAATGATTAAGAAAAAATA 1083

Db 50 GGGTTATCATCTCCAGATGCCAAATGATTAAGAAAAAATA 4

RESULT 10

LOCUS B1555512 829 bp mRNA linear EST 05-SEP-2001

DEFINITION 603239318F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292221 5', mRNA sequence.

ACCESSION B1555512

VERSION B1555512.1 GI:15442826

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 829)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA library preparation: Lothar Hennighausen, Inc.

DNA Sequencing by: Inocyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L14M11739 row: C column: 06

High quality sequence start: 2

High quality sequence stop: 753.

FEATURES

source

1..829

Location/Qualifiers

```

/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db.xref="taxon:10090"
/clone="IMAGE:5292221"
/clone_lib="NCI-CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; SalI;
Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT      209 a      214 c      216 g      190 t
ORIGIN
Query Match      13.3%; Score 513.6; DB 13; Length 829;
Best Local Similarity 83.1%; Pred. No. 6e-98;
Matches 680; Conservative 0; Mismatches 129; Indels 9; Gaps 8;
QY 5 GAGCTTCATGGAGTACAGGCGCGGCTCG-ACGGATTGCCATGGCCCGCTGCTG-GA 62
Db 1 GGGGCCCCATGGAGCCGCTGTGGGCGAAGCGCGAGTATGCGCGCTGTGTGA 60
QY 63 GTACGAG-CGACAGCTGGTGTGTAAGTCTGACACTGACGCGGCTAGTAGTGGCGCC 121
Db 61 GTACAGATGGGAGGAGGCTGTGTAAGTCTGACACTGACGCGGCTAGTAGTGGCGCC 120
QY 122 GCGGCGCTGGGCGGAGCC-GGCTCTCTTACCACTTTCTCCAGCTGCACCTGCCAGCC 180
Db 121 GCGGACTGGGAGCCGCTGTGCTTACCACTTTCTCCAGCTGCACCTGCCAGCC 180
QY 181 TGCCGATGCTGGTGTGCTCAACGACGCGCGGCGAGGAGATTTATCAATGACGTC 240
Db 181 TGCCGATGCTGGTGTGCTCAACGACGCGCGGCGAGGAGATTTATCAATGACGTC 240
QY 241 AAGATAGAAGAGTGAACACCTCCCTCGCGTGTACAAATGAATCACAACAACAGT 300
Db 241 AAGATAGAAGAGTGAACACCTCCCTCGCGAGTGTACAAATGAATGCAAGTAACAGT 300
QY 301 CGCTGTGAAGTTTACACACAGAGTGTGTATATTGGACAGTAGATTTGTTGTT 360
Db 301 CGCTGTGAAGTTTACACACAGAGTGTGTATATTGGACAGTAGATTTGTTGTT 360
QY 361 GACTTCTGATGATAGATACCTTCAATTTAATTACTGGCATCTTGTGTATAGAGCC 420
Db 361 GATTTCCTGACAGTATGATACCTTCAATTTAATTACTGGCATCTTGTGTATAGAGCC 420
QY 421 CACAGATTAATGCA-GTCTTGTCAAGAAGCATCTTGGCCCTTTCCGACAGAAAA 479
Db 421 CACAGATTAATGCA-GTCTTGTCAAGAAGCATCTTGGCCCTTTCCGACAGAAAA 480
QY 480 CAACGCTGTTTATTAAAGCTTTCACAGACAACTCTGTG-CCTTTGATACTGTTT 538
Db 481 CAAGGCGGGTTCATCAAGCTTTCACAGACAACTCTGTG-CCTTTGATACTGTTT 540
QY 539 GTCATGTGGAAGAAGTGTATGAATAATCTTTTGTGAGAACTGTATCTGTGCGCAAGT 598
Db 541 GTCAAGTGAAGAAGTGTATGCGAAGCTTTTGTGAGAGTCTTACTGTGCGCAAGT 600
QY 599 TCCATGTAGCAAGTAACTATTTTGAAGACAGCAAACTGAACTGTAGAAATCCATG 658
Db 601 TCCATGTAGCAAGTAACTATTTTGAAGACAGCAAACTGAACTGTAGAAATCCATG 660
QY 659 TTTTCATGACCTTACCTGCTGTATAGAG-ACCTGATATAGAG-ATTTAATG 715
Db 661 TGTTCATGACCTTACCTGCTGTATAGAG-ACCTGATATAGAG-ATTTAATG 720
QY 716 CATGCTTAAGAGAACTAAATGCCATAACCATGCTTGAAGTGAAGATTTATCTTAG 775
Db 721 CATGCTTAAGAGAACTAAATGTGCACACCGCTTCACTTGAAGTGAAGATTTATCTTAG 780

```

```

QY 776 AAAATGCTATGGAACCTTTTGACAAAGACAATCCGC 813
Db 781 AACAACTCTTGGGAAGCATTCGACAGACAAATCCGC 818

RESULT 11
LOCUS      BQ310815
DEFINITION OY-BN0046-150300-121-b11 BN0046 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BQ310815
VERSION    BQ310815.1 GI:20866223
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 515)
            Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bouidin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT    JOURNAL
            MEDLINE
            CONTACT: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 0:539-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/ICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?rl-QV3&t2-QV3-BN0046-
            150300-121-b11&t3=2000-03-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 515.
            Location/Qualifiers
                1..515
                    /organism="Homo sapiens"
                    /db.xref="taxon:9606"
                    /clone_lib="BN0046"
                    /dev_stage="Adult"
                    /note="Organ: breast_normal; Vector: puc18; Site:1: SmaI;
                    Site:2: SmaI; A mini-library was made by cloning products
                    derived from ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 - Ludwig Institute for Cancer Research)
                    profiles into the puc 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."
BASE COUNT      131 a      134 c      118 g      132 t
ORIGIN
Query Match      13.0%; Score 499.6; DB 14; Length 515;
Best Local Similarity 98.2%; Pred. No. 5.4e-95;
Matches 505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1944 AGAAGAGAGGACAGATGAAACAACCTTAGTAAAGGCGACAGCATCTGCAGA 2003
Db 2 AGAAGAGAGGAGCAGATGAAACAACCTTAGTAAAGGCGACAGCATCTGCAGA 61
QY 2004 TGTTCACCTGACACTCGGAAAGCCGTTGGCCAGGAGAGATGCTACACAGCAAGAT 2063
Db 62 TGTTCACCTGACACTCGGAAAGCCGTTGGCCAGGAGAGATGCTACACAGCAAGAT 121
QY 2064 AGTGTGATATGCGTGAATTTGGAAGGAGGCTTCATCTCGATCCATTCGCGGGGAT 2123
Db 122 AGTGTGATATGCGTGAATTTGGAAGGAGGCTTCATCTCGATCCATTCGCGGGGAT 181

```

QY 2124 TGACATTTGAACCCGTCAGTTAGAGGTGGAGATTACATCCTCAGTCACAGAAATGCGCT 2183
 |||||
 DB 182 TGACATTTGAACCCGTCAGTTAGAGGTGGAGATTACATCCTCAGTCACAGAAATGCGCT 241
 |||||
 QY 2184 GGAGCCCAAGAGATGATGATTTAAATCGGCTTTTAAATACGCCGCCCTTCACAGCCA 2243
 |||||
 DB 242 GGAGCCCAAGAGATGATGATTTAAATCGGCTTTTAAATACGCCGCCCTTCACAGCCA 301
 |||||
 QY 2244 GTGCATCTCCATCTCCCGCTACACAAAGCTCCCGTCTCTCATTTGAGTTCACCCCTAG 2303
 |||||
 DB 302 GTGCATCTCCATCTCCCGCTACACAAAGCTCCCGTCTCTCATTTGAGTTCACCCCTAG 361
 |||||
 QY 2304 CAAGCCCTTCTCTCTCAGTCTCCGAGGTGCTTTTTCAGAGATCTCCAGCAATGACAT 2363
 |||||
 DB 362 CAAGCCCTTCTCTCTCAGTCTCCGAGGTGCTTTTTCAGAGATCTCCAGCAATGACAT 421
 |||||
 QY 2364 TAGTCCCAAACTCAGTCTTCTTACACTTCCACAGACTACGGAATCTCTGTGCGCC 2423
 |||||
 DB 422 TAGTCCCAAACTCAGTCTTCTTACACTTCCACAGACTACGGAATCTCTGTGCGCC 481
 |||||
 QY 2424 CTCTCTCATGCAACGGCGGAGTTGTTGAGAG 2457
 |||||
 DB 482 CTCTCTCATGCAACGGCGGAGTTGTTGAGAG 515
 |||||

RESULT 12
 BM671280/c 524 bp mRNA 1linear EST 27-FEB-2002
 LOCUS
 DEFINITION
 UI-E-CQ1-agg-h-17-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
 ACCESSION
 BM671280
 VERSION
 BM671280.1 GI:18981178
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 524)
 AUTHORS
 Ronaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL
 Genome Res. 6 (9), 791-806 (1996)
 MEDLINE
 97044477
 COMMENT
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mscares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLY-A-tail.

FEATURES
 source
 Location/Qualifiers
 1..524
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-CQ1-agg-h-17-0-UI"
 /clone_lib="UI-E-CQ1"
 /tissue_type="optic nerve"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pTZ19-Pac (pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-CQ1 is a normalized cDNA library containing the
 following tissue(s): optic nerve. The library was
 constructed according to Ronaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pTZ19-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (GT)₁₈ tail. The sequence tag for this library is
 CCATTAATG. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI).
 TAG_LIB-UI-E-CQ1
 TAG_TISSUE-human optic nerve
 TAG_SEQ-CCATTAATG

BASE COUNT 119 a 136 c 127 g 142 t

ORIGIN

Query Match 12.9%; Score 499; DB 13; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7, 2e-95;
 Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TGGAGTCAGGGGACGCGGCTCGACGGATTGCCATGGCGCCGCTGCTGAGTACGAGGAC 73
 |||||
 DB 515 TGGAGTCAGGGGACGCGGCTCGACGGATTGCCATGGCGCCGCTGCTGAGTACGAGGAC 456
 |||||
 QY 74 AGCTGGTGTGGAAGCTGCTGACACCTGACGCGGCTAGTGTGCGCCGCGGCTCGGCG 133
 |||||
 DB 455 AGCTGGTGTGGAAGCTGCTGACACCTGACGCGGCTAGTGTGCGCCGCGGCTCGGCG 396
 |||||
 QY 134 CGGACCGGCTCTCTTACACCTTCTCCAGCTGCACTGCCACCGACCCCTGCTGCTGG 193
 |||||
 DB 395 CGGACCGGCTCTCTTACACCTTCTCCAGCTGCACTGCCACCGACCCCTGCTGCTGG 336
 |||||
 QY 194 TGTCTCAACACGACGCGCGCGGAGAGAGTATTATCAATCAGCTTAACATAGAGAG 253
 |||||
 DB 335 TGTCTCAACACGACGCGCGCGGAGAGAGTATTATCAATCAGCTTAACATAGAGAG 276
 |||||
 QY 254 TTGAACACCTCCCTCGCGCGGTGTAAACAATGAATCACAACGACAGCTGATGAAGTT 313
 |||||
 DB 275 TTGAACACCTCCCTCGCGCGGTGTAAACAATGAATCACAACGACAGCTGATGAAGTT 216
 |||||
 QY 314 ACACACAAGGTGTGTATTTTGGCACAAGTAGAGATCTTGTGACTTCTTGACTG 373
 |||||
 DB 215 ACACACAAGGTGTGTATTTTGGCACAAGTAGAGATCTTGTGACTTCTTGACTG 156
 |||||
 QY 374 ATGAAATACCTTCAGATTATTAATCTGCGATCTTGTGTATAGAGCCCAAGATATCG 433
 |||||
 DB 155 ATGAAATACCTTCAGATTATTAATCTGCGATCTTGTGTATAGAGCCCAAGATATCG 96
 |||||
 QY 434 AGCTTGTCAAGAAGCATTCATCTTGGCGCTCTTGGCCAGAAAACAACGCTGTTTGA 493
 |||||
 DB 95 AGCTTGTCAAGAAGCATTCATCTTGGCGCTCTTGGCCAGAAAACAACGCTGTTTGA 36
 |||||
 QY 494 TTAAGGCTTTCACAGACAA 512
 |||||
 DB 35 TTAAGGCTTTCACAGACAA 17
 |||||

RESULT 13
 BE818393/c 618 bp mRNA 1linear EST 21-SEP-2000
 LOCUS
 RCS-BN0282-010600-031-B08 BN0282 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 BE818393
 ACCESSION
 BE818393.1 GI:10250627
 VERSION
 BE818393.1
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 618)
 AUTHORS
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

TITLE
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
20202663

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=6t2-RC5-BN0282-010600-031-B08&t3=2000-06-01&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 427.

FEATURES

source
1. 618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0282"
/dev_stage="Adult"
/note="Organ: A mini-library was made by cloning products derived from ORSTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
Location/Qualifiers
132 a 155 c 114 g 217 t

BASE COUNT
ORIGIN

Query Match 12.6%; Score 485.6; DB 12; Length 618;
Best Local Similarity 99.0%; Pred. No. 4.9e-92;
Matches 499; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1221 TCCAGGTCACACTGATTTGTCAGAGATGACCGAATGTTCCAGCTGAGAGACTA 1280
DB 512 TTCAGGTCACACTGATTTGTCAGAGATGACCGAATGTTCCAGCTGAGAGACTA 453
QY 1281 TATCACTCTTGAGACGGA-GGCTTCTTATGAGGCTCTACAGAGAACTTTGAGAGG 1339
DB 452 TATCACTCTTGAGACGGA-GGCTTCTTATGAGGCTCTACAGAGAACTTTGAGAGG 393
QY 1340 ATAGCAAGCTGAGAGAGTGTGATGAATTTAGGAGAGAGACAGTTCAAGAGAAATTA 1399
DB 392 ATAGCAAGCTGAGAGAGTGTGATGAATTTAGGAGAGAGACAGTTCAAGAGAAATTA 333
QY 1400 GGAATCTCACAAGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1459
DB 332 GGAATCTCACAAGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 273
QY 1460 GAACCTCAAAAAAG 1519
DB 272 GAACCTCAAAAAAG 213
QY 1520 AACTGGAAGAGAGAGAGATGTCGAGAGAGAGATATCGTGAAGAAATTAAGCAGTACCAG 1579
DB 212 AACTGGAAGAGAGAGAGATGTCGAGAGAGAGATATCGTGAAGAAATTAAGCAGTACCAG 153
QY 1580 AAACCTCCCGGAG 1639
DB 152 AAACCTCCCGGAG 93
QY 1640 CTTTCGGAATCTGGAAG 1699
DB 92 CTTTCGGAATCTGGAAG 33

QY 1700 CCTATGCTCTGACAGGACTACTAC 1723
DB 32 CCTATGCTCTGACAGGACTACTAC 9

RESULT 14
B0944411
LOCUS
DEFINITION
IMAGE:6440053 5', mRNA sequence.
ACCESSION
B0944411
VERSION
KEYWORDS
SOURCE
house mouse.
ORGANISM
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs@emil.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM13961 row: a column: 14
High quality sequence stop: 536.
Location/Qualifiers
1. 923
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6440053"
/clone_lib="NCI CGAP Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

FEATURES

source
1. 923
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6440053"
/clone_lib="NCI CGAP Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT
ORIGIN

Query Match 12.5%; Score 482.2; DB 14; Length 923;
Best Local Similarity 78.7%; Pred. No. 2.6e-91;
Matches 588; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1823 GGGTTTACTTCTTATATACGAGGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1882
DB 17 GGGTTTACTTCTTATATACGAGGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 76
QY 1883 GGAAGAAGAGAGAGCTTTGAAAAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1942
DB 77 GGAAGAAGAGAGAGCTTTGAAAAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 136
QY 1943 AAGAAG 2002
DB 137 AAGAAG 196
QY 2003 ATGTTTCCACTGACCTCGGAAGCCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAG 2062
DB 197 ATGTTTCCACTGACCTCGGAAGCCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAG 256
QY 2063 TAGTTTGGAGATATCGGAG 2122
DB 257 TAGTTTGGAGATATCGGAG 316

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:43:29 : Search time 96.2334 Seconds
(without alignments)
12291.496 Million cell updates/sec

Title: US-09-919-497-14

Perfect score: 3857
Sequence: 1 ggaagagcttcacatgagagtc.....agaaatacaactaagtc 3857

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2-6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2-6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2-6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2-6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2-6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2-6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2718	70.5	2718	4	US-09-651-656-14
2	2718	70.5	2718	4	US-09-650-855-14
3	1101	28.5	1101	4	US-09-651-656-12
4	1101	28.5	1101	4	US-09-650-855-12
5	78	2.0	7218	1	US-08-237-463-14
6	68.6	1.8	80246	4	US-09-078-294-4
7	68.2	1.8	162450	4	US-09-345-882-1
8	68	1.8	45716	4	US-08-965-048-5
9	68	1.8	45989	4	US-08-965-048-6
10	67.8	1.8	168575	4	US-09-426-290-1
11	67.2	1.7	80595	4	US-09-078-294-3
12	67	1.7	176373	3	US-09-128-155-17
13	66.6	1.7	84495	4	US-09-797-906-3
14	66.4	1.7	45716	4	US-08-965-048-5
15	66.4	1.7	45989	4	US-08-965-048-6
16	66.2	1.7	14796	4	US-08-975-080-35
17	66.2	1.7	14796	4	US-09-630-706-10
18	66.2	1.7	14796	4	US-09-496-694B-3
19	66.2	1.7	55827	4	US-09-813-133A-3
20	66.2	1.7	70000	4	US-09-851-896-3
21	66	1.7	6792	4	US-09-374-454-20
22	66	1.7	8224	2	US-09-010-398-14
23	66	1.7	8224	2	US-09-366-260-14
24	66	1.7	246240	2	US-08-724-394A-20
25	66	1.7	246240	2	US-08-724-394A-21
26	66	1.7	246240	2	US-08-724-394A-22
27	65.8	1.7	35100	1	US-08-306-691B-19

c	28	65.8	1.7	35100	5	PCT-US93-06251-19	Sequence 19, Appl
c	29	65.8	1.7	40352	3	US-08-846-111D-15	Sequence 15, Appl
c	30	65.8	1.7	40352	4	US-09-443-077-15	Sequence 15, Appl
c	31	65.8	1.7	50000	4	US-09-146-053-4	Sequence 4, Appl1
c	32	65.6	1.7	29629	4	US-09-729-995-3	Sequence 3, Appl1
c	33	65.6	1.7	43795	4	US-08-742-185-101	Sequence 10, App
c	34	65.4	1.7	99500	4	US-09-798-096-10	Sequence 10, Appl
c	35	65	1.7	11811	4	US-09-078-294-7	Sequence 7, Appl1
c	36	64.8	1.7	2907	2	US-09-018-628-17	Sequence 17, Appl
c	37	64.8	1.7	2907	3	US-09-273-378-17	Sequence 17, Appl
c	38	64.8	1.7	2907	4	US-09-018-635-26	Sequence 26, Appl
c	39	64.8	1.7	2907	4	US-09-467-642-3	Sequence 3, Appl1
c	40	64.8	1.7	16389	4	US-09-741-154-3	Sequence 3, Appl1
c	41	64.6	1.7	1001	4	US-09-641-638-354	Sequence 354, App
c	42	64.6	1.7	2908	3	US-08-487-789-1	Sequence 1, Appl1
c	43	64.6	1.7	14581	4	US-08-520-373D-4	Sequence 4, Appl1
c	44	64.6	1.7	14636	4	US-09-173-914-6	Sequence 6, Appl1
c	45	64.4	1.7	3366	4	US-09-345-650-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1									
US-09-651-656-14									
Sequence 14, Application US/09651656									
Patent No. 6340566									
GENERAL INFORMATION:									
APPLICANT: MCCUTHEN-MALONEY, SANDRA									
TITLE OF INVENTION: LAWRENCE LIVERMORE NATIONAL LABORATORY									
TITLE OF INVENTION: DETECTION AND QUANTIFICATION OF SINGLE NUCLEOTIDE									
TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,									
TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES									
FILE REFERENCE: IL-10689									
CURRENT APPLICATION NUMBER: US/09/651,656									
CURRENT FILING DATE: 2000-08-29									
PRIOR APPLICATION NUMBER: 60/192,764									
PRIOR FILING DATE: 2000-03-28									
NUMBER OF SEQ ID NOS: 106									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 14									
LENGTH: 2718									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-651-656-14									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
OY	46	ATGGCGCGCTCTGAGTACGACGACGCTGTGTTGAACCTCTCGACACTGACGG	105						
DB	1	ATGGCGCGCTCTGAGTACGACGACGACGCTGTGTTGAACCTCTCGACACTGACGG	60						
OY	106	CTAGAGTGTGCGCGCGCGCGCTGCGCGCGCGCTCTTACACTTCTCCAGCTG	165						
DB	61	CTAGAGTGTGCGCGCGCGCGCTGCGCGCGCGCTCTTACACTTCTCCAGCTG	120						
OY	166	CAGTCCACCCAGCTGCTGTGCTGTGCTCAACACGACCGCGCGAGGAGATAT	225						
DB	121	CAGTCCACCCAGCTGCTGTGCTGTGCTCAACACGACCGCGCGAGGAGATAT	180						
OY	226	TTTATCAATCAGCTGAGATAGAGAGAGTGAACCTCTCGCGGTAAACAATGAA	285						
DB	181	TTTATCAATCAGCTGAGATAGAGAGAGTGAACCTCTCGCGGTAAACAATGAA	240						
OY	286	ATCAACAACACAGCTGCTATGATTAACACACAGAGTGTATTTTCCGACAGT	345						
DB	241	ATCAACAACACAGCTGCTATGATTAACACACAGAGTGTATTTTCCGACAGT	300						
OY	346	AGGATCTTGTGTTGACTTCTTGACTGATAGAAATACCTTGAGATTATTACTGCATC	405						
DB	301	AGGATCTTGTGTTGACTTCTTGACTGATAGAAATACCTTGAGATTATTACTGCATC	360						

QY	406	TTGTGTATAGAGCCCAACAGATATATTCAGCTGTGTCCAAAGAGCATTCATCTTGGCCTC	465
Db	361	TTGGGTATATAGGCCCAACAGATATATTCAGATCTGTCTCAAGAAAGCATTCATCTTGGCCTC	420
QY	466	TTTGGCCAGAAAACCAACGCGTTTTATTTAAAGCTTTTACAGACAAATCGTGGCCCTT	525
Db	421	TTTGGCCAGAAAACCAACGCTGTTTTATTAAAGCTTTTACAGACAAATCTGTGCCCTT	480
QY	536	GATACGTGTTTTTGTCACTGTGGAAAGAGTATGAGAAATCTTTTGTGAGAAACTGTAT	585
Db	481	GATACGTGTTTTTGTCACTGTGGAAAGAGTATGAGAAATCTTTTGTGAGAAACTGTAT	540
QY	586	CTGTGGCCCAAGTTCATGTATAGCACTTAATCTATTTTGAACAGACAAACCTTAAGT	645
Db	541	CTGTGGCCCAAGTTCATGTATAGCACTTAATCTATTTTGAACAGACAAACCTTAAGT	600
QY	646	GTAGAAATCCATGTGTTTGTATGACACCTTACCTCTGCTATACAGACTCTTACTGTGAC	705
Db	601	GTAGAAATCCATGTGTTTGTATGACACCTTACCTCTGCTATATACAGACTCTTACTGTGAC	660
QY	706	ATTTTAAATGCATGTCTTAAAGGAATTAATGCCATTAACCATCGCTTGAAGTGAAGAT	765
Db	661	ATTTTAAATGCATGTCTTAAAGGAATTAATGCCATTAACCATCGCTTGAAGTGAAGAT	720
QY	766	TTATCTTTAGAAATTCGATTTGGAAAACTTTTGAACAACATCCGCCATTAATGTGAT	825
Db	721	TTATCTTTAGAAATTCGATTTGGAAAACTTTTGAACAACATCCGCCATTAATGTGAT	780
QY	826	CCTTTGTGGCCACAGCTTGGAGCCCAAGACTAAATCCCTTACAGATTTGAAGATATTA	885
Db	781	CCTTTGTGGCCACAGCTTGGAGCCCAAGACTAAATCCCTTACAGATTTGAAGATATTA	840
QY	886	CGAAGTTTGGCGAGTATCTCTCGAGTATGATGTGTGCATATTCATTCATTCCTGTGAA	945
Db	841	CGAAGTTTGGCGAGTATCTCTCGAGTATGATGTGTGCATATTCATTCATTCCTGTGAA	900
QY	946	TCTCTGAGACCAACGGAAGAAAAAGCTTTTGGTGCAGAAATTCAGGTTGGCTGTCTTACTCC	1005
Db	901	TCTCTGAGACCAACGGAAGAAAAAGCTTTTGGTGCAGAAATTCAGGTTGGCTGTCTTACTCC	960
QY	1006	AGCACTCTGATGTTTATTAATTCGTCGAGCAAGGGTTATTCATCTTCCAGATCCCAAATG	1065
Db	961	AGCACTCTGATGTTTATTAATTCGTCGAGCAAGGGTTATTCATCTTCCAGATCCCAAATG	1020
QY	1066	AGTAAAAAGAAAAATATCTGAAAAAATGSAATTTAAAGAGGGGAGAAACAAAAAAAG	1125
Db	1021	AGTAAAAAGAAAAATATCTGAAAAAATGSAATTTAAAGAGGGGAGAAACAAAAAAAG	1080
QY	1126	GACCTGTCTTATGAAACCAACCAAGTGGAGGCCTGACCTGSAATTTAAAGAAAT	1185
Db	1081	GACCTGTCTTATGAAACCAACCAAGTGGAGGCCTGACCTGSAATTTAAAGAAAT	1140
QY	1186	GAGGAGAAATTAAGAGAGTGAAGCTCTTGGTGTGCACAGCTCAAGTACTGATTTGTGCA	1245
Db	1141	GAGGAGAAATTAAGAGAGTGAAGCTCTTGGTGTGCACAGCTCAAGTACTGATTTGTGCA	1200
QY	1246	AGTGATGACCGACATCTTCCCAAGCTGAGAGACTATATCACTCTTGGAGCGGAGCCTTC	1305
Db	1201	AGTGATGACCGACATCTTCCCAAGCTGAGAGACTATATCACTCTTGGAGCGGAGCCTTC	1260
QY	1306	TTATTTAGAGGCTCTACAGGAACCTTTGAAAGAGATAGCAAAAGCTGAAGATCTGTGATG	1365
Db	1261	TTATTTAGAGGCTCTACAGGAACCTTTGAAAGAGATAGCAAAAGCTGAAGATCTGTGATG	1320
QY	1366	AAATTTAGGAAGAAGACAGTTCGAAGAGAAATTTAGCAAAATCTCACAAAAAGACTTAAGAC	1425
Db	1321	AAATTTAGGAAGAAGACAGTTCGAAGAGAAATTTAGCAAAATCTCACAAAAAGACTTAAGAC	1380
QY	1426	CCCCCAAAACAAGAGCGGCTTCTACCAAGAAAGAACCTTCAAAAGAAAAACGGAAG	1485
Db	1381	CCCCCAAAACAAGAGCGGCTTCTACCAAGAAAGAACCTTCAAAAGAAAAACGGAAG	1440

QY	1486	TTGACCTTA	ACTCA	AAATG	TGTA	AGCAAA	ACTG	GAAGAA	CTGGA	ACACG	ACGAC	ATGTC	AG	1545													
DB	1441	TTGACCTTA	ACTCA	AAATG	TGTA	AGCAAA	CTGGA	ACAA	CTGGA	ACAA	CTGGA	ATGTC	AG	1500													
QY	1546	GAAGATAT	TGCTCG	GAAGAA	TAAT	TAG	CAGTA	AGCCG	GAAGAA	CTGCC	CCG	AA	TAAT	TAGAT	1605												
DB	1501	GAAGATAT	TGCTCG	GAAGAA	TAAT	TAG	CAGTA	AGCCG	GAAGAA	CTGCC	CCG	AA	TAAT	TAGAT	1560												
QY	1606	GAAGAAT	TTGAT	GTAAAT	TTTG	TG	CAT	CGGAT	GTGCT	TTG	CGA	ATGCT	TA	GAAGAC	CTC	1665											
DB	1561	GAAGAAT	TTGAT	GTAAAT	TTTG	TG	CAT	CGGAT	GTGCT	TTG	CGA	ATGCT	TA	GAAGAC	CTC	1620											
QY	1666	ACTATAC	CAATC	CGCG	CTTG	GGGT	TG	GAG	GACCC	CTAT	ATG	CTG	CA	CAAGG	TA	ACTAT	1725										
DB	1621	ACTATAC	CAATC	CGCG	CTTG	GGGT	TG	GAG	GACCC	CTAT	ATG	CTG	CA	CAAGG	TA	ACTAT	1680										
QY	1726	GAAGTGA	GCC	AGATAC	GTGG	TTCT	TA	TAG	CGAC	AGAC	CTTA	CTG	TTG	TG	CGAC	CTT	1785										
DB	1681	GAAGTGA	GCC	AGATAC	GTGG	TTCT	TA	TAG	CGAC	AGAC	CTTA	CTG	TTG	TG	CGAC	CTT	1740										
QY	1786	GAATTTA	CAG	GGG	GAGT	AGG	CC	TGG	GA	AA	CTC	TG	AGG	TTT	ACT	TTCT	TA	TAGGA	1845								
DB	1741	GAATTTA	CAG	GGG	GAGT	AGG	CC	TGG	GA	AA	CTC	TG	AGG	TTT	ACT	TTCT	TA	TAGGA	1800								
QY	1846	GTTTCA	ACT	GA	TGA	AGAA	CA	ACC	CTA	ATC	TAC	TG	CTT	CCG	GA	AA	CA	AA	AG	CA	AG	CTTT	TGA	1905			
DB	1801	GTTTCA	ACT	GA	TGA	AGAA	CA	ACC	CTA	ATC	TAC	TG	CTT	CCG	GA	AA	CA	AA	AG	CA	AG	CTTT	TGA	1860			
QY	1906	AAACTCA	T	AA	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1965		
DB	1861	AAACTCA	T	AA	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1920		
QY	1966	ACAATCT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	2025		
DB	1921	ACAATCT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	GA	ACT	TA	1980		
QY	2026	GCCG	TG	CC	AG	GA	CA	AG	AA	TG	TAC	AC	AG	CA	AA	AG	CA	T	AG	TTG	TG	CG	TA	AT	2085		
DB	1981	GCCG	TG	CC	AG	GA	CA	AG	AA	TG	TAC	AC	AG	CA	AA	AG	CA	T	AG	TTG	TG	CG	TA	AT	2040		
QY	2086	CGAATG	AG	CTT	CA	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	2145		
DB	2041	CGAATG	AG	CTT	CA	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	2100		
QY	2146	GAGGT	TG	GA	G	A	T	T	C	T	C	A	A	A	A	T	T	C	G	G	T	G	A	A	T	2205	
DB	2101	GAGGT	TG	GA	G	A	T	T	C	T	C	A	A	A	A	T	T	C	G	G	T	G	A	A	T	2160	
QY	2206	TTAAT	CG	CG	CTT	T	T	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2265	
DB	2161	TTAAT	CG	CG	CTT	T	T	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2220	
QY	2266	TACA	AG	CG	T	CC	G	T	CTT	CT	G	A	T	T	G	A	T	T	G	A	C	C	T	A	C	T	2325
DB	2221	TACA	AG	CG	T	CC	G	T	CTT	CT	G																

|||||
Db 2521 TTGTTAAATATCCAGGGGTGATGCCAAAACCTCCGCTCTTATGACACACCTTAAG 2580
OY 2626 AACATCGAGATTAAGACAGCCCTGTACAAAGAGAGCTCAGAGTATCTGGGAGATGT 2685
Db 2581 AACATCGAGATTAAGACAGCCCTGTACAAAGAGAGCTCAGAGTATCTGGGAGATGT 2640
OY 2686 GCAAAATGCCAAACAGCTTTATGATTTCAATTCACACCTCTTTTGGAGAGTGTATCAAA 2745
Db 2641 GCAAAATGCCAAACAGCTTTATGATTTCAATTCACACCTCTTTTGGAGAGTGTATCAAA 2700
OY 2746 GGAAGAGGAAAAAGTGA 2763
Db 2701 GGAAGAGGAAAAAGTGA 2718

RESULT 2
US-09-650-855-14
; Sequence 14, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-650-855-14

Query Match 70.5%; Score 2718; DB 4; Length 2718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 421 TTTCGCCAAGAAAAACAAACGTGTTTATTAAGCTTTCACAGACATCTGTGCTTT 480
OY 526 GATACGGTTTTTGTGATGTGAAGAGATGAGAAATCTTTTGTAGGAACTGTAT 585
Db 481 GATACGGTTTTTGTGATGTGAAGAGATGAGAAATCTTTTGTAGGAACTGTAT 540
OY 586 CTGTGGCCAAAGTTCCATGTAGACAGTAAACTCATTTTTTGAACAGACAAACCTGAAGT 645
Db 541 CTGTGGCCAAAGTTCCATGTAGACAGTAAACTCATTTTTTGAACAGACAAACCTGAAGT 600
OY 646 GTAGAAATCCATGTTTCTATGACACCTTACACTGTTGCTATACAGACTGTATCTGAC 705
Db 601 GTAGAAATCCATGTTTCTATGACACCTTACACTGTTGCTATACAGACTGTATCTGAC 660
OY 706 ATTTAAATGCATGTCTAAGAACTAAATGCGAATCCCATGCTGTGAAGTGAAGT 765
Db 661 ATTTAAATGCATGTCTAAGAACTAAATGCGAATCCCATGCTGTGAAGTGAAGT 720
OY 766 TTATCTTTAGAAAATGCTATTTGAAAACCTTTTGAACAAGACAAATCCGCAATATCTGAT 825
Db 721 TTATCTTTAGAAAATGCTATTTGAAAACCTTTTGAACAAGACAAATCCGCAATATCTGAT 780
OY 826 CCTTTGTGGCACACCTTGTGAGCCAAAGCTAAATCTTATGTTAGATTTGAATATTA 885
Db 781 CCTTTGTGGCACACCTTGTGAGCCAAAGCTAAATCTTATGTTAGATTTGAATATTA 840
OY 886 GCAACTTGTGCACTATCTCTCAGATGATGATGTCACATTTCTATCTTCTGGA 945
Db 841 GCAACTTGTGCACTATCTCTCAGATGATGATGTCACATTTCTATCTTCTGGA 900
OY 946 TCTCTGAGAGCAACGAAAAAGCTTTTGTGCAATTTAGTTGAGTCTTTCTTCTGACTCC 1005
Db 901 TCTCTGAGAGCAACGAAAAAGCTTTTGTGCAATTTAGTTGAGTCTTTCTTCTGACTCC 960
OY 1006 ACCACCTGATGTTTATTAATGCTGAGCAAGGTTTATCATCTTCCAGATCCAAATG 1065
Db 961 ACCACCTGATGTTTATTAATGCTGAGCAAGGTTTATCATCTTCCAGATCCAAATG 1020
OY 1066 AGTAAAAAGAAAAATATCTGAATAATGAAATTAAGAAAGGGAACAAACAAAG 1125
Db 1021 AGTAAAAAGAAAAATATCTGAATAATGAAATTAAGAAAGGGAACAAACAAAG 1080
OY 1126 GAACTGTCTCTAAGAAAGAACCCAAAGTGGAGCACTGACTGAGATTTAAAGAAAT 1185
Db 1081 GAACGTGCTCTAAGAAAGAACCCAAAGTGGAGCACTGACTGAGATTTAAAGAAAT 1140
OY 1186 GAGCGAAAAATTAAGAGAGTGAAGCTTGTGTGTCAGAGTCAAGTACTGATTTGTGCA 1245
Db 1141 GAGCGAAAAATTAAGAGAGTGAAGCTTGTGTGTCAGAGTCAAGTACTGATTTGTGCA 1200
OY 1246 AGTATGACCCGAACATGTTTCCAGCTGAGACATATATCAGTCTTGGAGCGAGGCTTC 1305
Db 1201 AGTATGACCCGAACATGTTTCCAGCTGAGACATATATCAGTCTTGGAGCGAGGCTTC 1260
OY 1306 TTATGAGGCTCTTACAGAAAACTTTGAGAAAGATATCAAGCTGAGAAAGTGTGATG 1365
Db 1261 TTATGAGGCTCTTACAGAAAACTTTGAGAAAGATATCAAGCTGAGAAAGTGTGATG 1320
OY 1366 AAATTTAGGAAGAGACAGTTCAAAGAGATTTAGAAATCTCACAAAGACCTTAAGAC 1425
Db 1321 AAATTTAGGAAGAGACAGTTCAAAGAGATTTAGAAATCTCACAAAGACCTTAAGAC 1380
OY 1426 CCCCCAAAACAAAGAGGCGTCTTCCAAAGAAAGAAACCCCTCAAAAGAAAAACGGAG 1485
Db 1381 CCCCCAAAACAAAGAGGCGTCTTCCAAAGAAAGAAACCCCTCAAAAGAAAAACGGAG 1440
OY 1486 TTGACCTTAACCTCAAAATGTGAGAAACCTGAAAGACTGGAAGAGAGATGTGAG 1545
Db 1441 TTGACCTTAACCTCAAAATGTGAGAAACCTGAAAGACTGGAAGAGAGATGTGAG 1500
OY 1546 GAAAGATATCTGTCGAGAAATTAAGCAGTACCCAGAAACCTGCCGGAATTAAGCAT 1605
Db 1501 GAAAGATATCTGTCGAGAAATTAAGCAGTACCCAGAAACCTGCCGGAATTAAGCAT 1560

|||||
Db 541 CTGTGGCCAGAGTTCATGAGTAGCACTAACTCATTTTATAGAACAGCAAAACCTGAAGTT 600
Qy 646 GTAGAAATCCATGTTTCTATGACACCTTACCATGCTTGTATACAGACTGCTATAGTGCAC 705
Db 601 GTAGAAATCCATGTTTCTATGACACCTTACCATGCTTGTATACAGACTGCTATAGTGCAC 660
Qy 706 ATTTAAATGATGCTTAAAGAACTAAATGCAATACCATGCTTGAAGTGAAGAT 765
Db 661 ATTTAAATGATGCTTAAAGAACTAAATGCAATACCATGCTTGAAGTGAAGAT 720
Qy 766 TTATCTTTAGAAAATGCTATTGAAAACCTTTTGACAAGACATCCGCAATATCTGAT 825
Db 721 TTATCTTTAGAAAATGCTATTGAAAACCTTTTGACAAGACATCCGCAATATCTGAT 780
Qy 826 CCTTTGTGGCAGCAGCTTGGAGCCAGACACTAAATCCTTAGTTCAGAGATTGAAGATTA 885
Db 781 CCTTTGTGGCAGCAGCTTGGAGCCAGACACTAAATCCTTAGTTCAGAGATTGAAGATTA 840
Qy 886 CGAAGCTTGTGAGAGTATCTCTCAGATGATGTTGTACATTTCTTAATCTTCTGGA 945
Db 841 CGAAGCTTGTGAGAGTATCTCTCAGATGATGTTGTACATTTCTTAATCTTCTGGA 900
Qy 946 TCTCTGAGAGCAAGGAAAAGCTTTTGTGCAATTCAGGTTGGCTTTCTTGACTCC 1005
Db 901 TCTCTGAGAGCAAGGAAAAGCTTTTGTGCAATTCAGGTTGGCTTTCTTGACTCC 960
Qy 1006 AGCAGCTTCATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAAATG 1065
Db 961 AGCAGCTTCATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAAATG 1020
Qy 1066 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAAGGGAAGAAAACAAAAG 1125
Db 1021 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAAGGGAAGAAAACAAAAG 1080
Qy 1126 GAAGTGTCTCTAGAAAGCAAC 1146
Db 1081 GAAGTGTCTCTAGAAAGCAAC 1101

RESULT 4
US-09-650-855-12
; Sequence 12, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650/855
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-650-855-12

Query Match 28 5%; Score 1101; DB 4; Length 1101;
Best Local Similarity 100.0%; Prid. No. 1.9e-284;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 ATGGCGCGCTGCTGAGTACAGCAGCTGTGCTGAGACTGCTGAGACTGACGG 105
Db 1 ATGGCGCGCTGCTGAGTACAGCAGCTGTGCTGAGACTGCTGAGACTGACGG 60
Qy 106 CTAGTAGTGTGCGCGCGGCTCGCGCGGAGACCGGCTCTCTACCACTTTCTCCAGCTG 165
|||||

Db 61 CTAGTAGTGTGCGCGCGGCTCGCGCGGAGACCGGCTCTCTACCACTTTCTCCAGCTG 120
Qy 166 CACTGCCACCCAGCCTGCTGCTGCTGCTGCTCAACACAGCAGCGCGGAGAGAGAT 225
Db 121 CACTGCCACCCAGCCTGCTGCTGCTGCTGCTCAACACAGCAGCGCGGAGAGAGAT 180
Qy 226 TTATCATCATGAGCTGAAGATGAAGAGTTGAACACCTCCCTCGCGGTGAACAAATGA 285
Db 181 TTATCATCATGAGCTGAAGATGAAGAGTTGAACACCTCCCTCGCGGTGAACAAATGA 240
Qy 286 ATCAACAGCAACAGTGCCTATGAAGTTTACACACAGAGTGTGTATATTTCGCAAGT 345
Db 241 ATCAACAGCAACAGTGCCTATGAAGTTTACACACAGAGTGTGTATATTTCGCAAGT 300
Qy 346 AGGATACCTTTGTGCTTCTTCTGACTATGAATATACCTTCAGATTAACTGCAATC 405
Db 301 AGGATACCTTTGTGCTTCTTCTGACTATGAATATACCTTCAGATTAACTGCAATC 360
Qy 406 TTGTGTATAGAGCCACAGAAATATGAGTGTGTGCAAGAAAGCATCTTGGCGCTC 465
Db 361 TTGTGTATAGAGCCACAGAAATATGAGTGTGTGCAAGAAAGCATCTTGGCGCTC 420
Qy 466 TTTCCGACAGAAAACAAACGTGTTTATTAAGCTTTACAGACAAATGCTGCTTT 525
Db 421 TTTCCGACAGAAAACAAACGTGTTTATTAAGCTTTACAGACAAATGCTGCTTT 480
Qy 526 GATACCTGTTTTTGTGATGTGGAAGAGTATGAGAAATCTTTTGTGAGAACTGTAT 585
Db 481 GATACCTGTTTTTGTGATGTGGAAGAGTATGAGAAATCTTTTGTGAGAACTGTAT 540
Qy 586 CTGTGCGCAGGTTTCATGTGACATTAATGCTATTTTGAAGACAGCAAAACCTGAAGT 645
Db 541 CTGTGCGCAGGTTTCATGTGACATTAATGCTATTTTGAAGACAGCAAAACCTGAAGT 600
Qy 646 GTAGAAATCCATGTTTCTATGACACCTTACCATGCTTGTATACAGACTGCTATAGTGCAC 705
Db 601 GTAGAAATCCATGTTTCTATGACACCTTACCATGCTTGTATACAGACTGCTATAGTGCAC 660
Qy 706 ATTTAAATGATGCTTAAAGAACTAAATGCAATACCATGCTTGAAGTGAAGAT 765
Db 661 ATTTAAATGATGCTTAAAGAACTAAATGCAATACCATGCTTGAAGTGAAGAT 720
Qy 766 TTATCTTTAGAAAATGCTATTGAAAACCTTTTGACAAGACATCCGCAATATCTGAT 825
Db 721 TTATCTTTAGAAAATGCTATTGAAAACCTTTTGACAAGACATCCGCAATATCTGAT 780
Qy 826 CCTTTGTGGCAGCAGCTTGGAGCCAGACACTAAATCCTTAGTTCAGAGATTGAAGATTA 885
Db 781 CCTTTGTGGCAGCAGCTTGGAGCCAGACACTAAATCCTTAGTTCAGAGATTGAAGATTA 840
Qy 886 CGAAGCTTGTGAGAGTATCTCTCAGATGATGTTGTACATTTCTTAATCTTCTGGA 945
Db 841 CGAAGCTTGTGAGAGTATCTCTCAGATGATGTTGTACATTTCTTAATCTTCTGGA 900
Qy 946 TCTCTGAGAGCAAGGAAAAGCTTTTGTGCAATTCAGGTTGGCTTTCTTGACTCC 1005
Db 901 TCTCTGAGAGCAAGGAAAAGCTTTTGTGCAATTCAGGTTGGCTTTCTTGACTCC 960
Qy 1006 AGCAGCTTCATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAAATG 1065
Db 961 AGCAGCTTCATGTTTATTAATGCTCGAGCAAGGTTTATCATCTTCCAGATGCCAAAATG 1020
Qy 1066 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAAGGGAAGAAAACAAAAG 1125
Db 1021 AGTAAAAAGAAAAATATCTGAAAAATGGAATTAAGAAGGGAAGAAAACAAAAG 1080
Qy 1126 GAAGTGTCTCTAGAAAGCAAC 1146
Db 1081 GAAGTGTCTCTAGAAAGCAAC 1101

RESULT 5
US-08-232-463-14/c

```
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; JTELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F15
; US-08-232-463-14

Query Match          2.0%; Score 78; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 1,4e-10;
Matches 15; Conservative 242; Mismatches 137; Indels 0; Gaps 0;

QY 1035 AAGGTTTATCATCTCCAGATGCCAAATGAGTAAAGAAAGAAATATCTGAAATAAT 1094
    ||| ||| |
DB 1447 AACAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1388
QY 1095 GGAATTAAGAAGGAGGAGCAAAAGAAAGAACTGCTCTGAGAAAGCAACCAAGTG 1154
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1387 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1328
QY 1155 GGAGCAGCTGAGTGAAGTATTAAGAAATTTGAGCAGAAATTAAGAGAGTGAAGCTCT 1214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1327 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1268
QY 1215 TGCTGTCGACGTCAGTACTGTTGTGCAAGTGAAGCCGAACGTCTCCAGCTGAG 1274
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1267 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1208
QY 1275 AGACTATATCACTCTTGAGCGAGGCTCTTATTGAGGCTCTAGAGAAACCTTTGA 1334
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1207 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1148
QY 1335 GAAGATAGCAAGTGAAGAGCTGGATGAATTTAGAGAGGAAGAGCTTCAAGAG 1394
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 1147 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1088
QY 1395 AATTAGAAATCTCAAAAGACCTAAAGACCC 1428
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1087 RRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCC 1054

RESULT 6
US-09-078-294-4/C
; Sequence 4, Application US/09078294
; Patent No. 6255211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contlig
US-09-078-294-4

Query Match          1.8%; Score 68.6; DB 4; Length 80246;
Best Local Similarity 70.2%; Pred. No. 1.7e-07;
Matches 92; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 3723 CAGCATCATATCTCTAGAAATACCTAATTGGAGTGAGCCGAGATCGACACACTGC 3782
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4751 CAGGAGATCATCTTGAACCCAGGAGCGAAGTTGAGTGAGCGGATCGCCACTAC 4692
QY 3783 ACCCTGCTGGCGGACAGAGTGAAGTGTCTTATTACAAAAGAAAGAAAGAA 3842
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4691 ACTCCAGCCCTGGGTGACAGAGCACTGTCTAATAAAAAAAAAAAAAAAAAAAAA 4632
QY 3843 TACACCTAAG 3853
DB 4631 AAAAAAAAAAAG 4621

RESULT 7
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399375
; GENERAL INFORMATION:
; APPLICANT: Bouguetleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-
; FILE REFERENCE: GENSET 031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
```


FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:

NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele

```

LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

```

Query Match	1.8%	Score	68.2	DB	4	Length	162450
Best Local	Similarly	Pred	No. 3.2e-07				
Matches	62	Conservative	0	Mismatches	23	Indels	0
						Gaps	0

Oy 3748 ACCTAATTGGCAGTGCAGCCGAGANTCGCACCACTGCACCCCTGCCCTTG66CGACAGATGTAG 3807

Db 99615 ACGGAGCTTGCAGTCGACGCCAAGATGCGCGCACTGCACCTCCAGCCTGGCGGACAGAGCCAG 9967

QY	3808	ACTTGTCTCTATTACAAAAAGAAAAGAAATACACCTAA	3852
Db	99675	ACTCTGTCTCAAAAAAAAAAAGAGAACAAAAACAAAAAA	99719

RESULT 8
 US-08-965-048-5
 Sequence 5, Application US/08965048
 Patent No. 6323244
 GENERAL INFORMATION:
 APPLICANT: Chen, Hong
 APPLICANT: Frelmer, Nelson
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 FILE REFERENCE: 7853-093
 CURRENT APPLICATION NUMBER: US/08/965,048
 CURRENT FILING DATE: 1997-11-05
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 45716
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-08-965-048-5

Query Match	1.8%	Score 68	DB 4	Length 45716
Best Local Similarly	83.7%	Pred. No. 1.8e-07		
Matches 77	Conservative	0	Mismatches 15	Indels 0
				Gaps 0

Qy	3755	TGGCAGTGTGAGCCGACGATGCGACCACTGCAACCCCTGCCCTGGCGACAGAGTGTGT	3814
Db	21289	TTGCAGTGTGAGCCGACGATGCGACCACTGCAACCCCTGGCGACAGAGTGTGT	21348

Qy	3815	CTCTATTACAAAAAGAAAAGAAATACA	3846
Db	21349	CTCAAAACAAAAAAGAAAAAAGAAAAAGA	21380

```

RESULT 9
US-08-965-048-6
: Sequence 6, Application US/08965048
: Patent No. 6323244
: GENERAL INFORMATION:
: APPLICANT: Chen, Hong
: APPLICANT: Frelmer, Nelson
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
: TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
: FILE REFERENCE: 7853-093
: CURRENT APPLICATION NUMBER: US/08/965,048
: CURRENT FILING DATE: 1997-11-05
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 6
: LENGTH: 45989
: TYPE: DNA
: ORGANISM: Homo sapiens
US-08-965-048-6

```

Query Match	1.8%	Score 68	DB 4	Length 45989
Best Local Similarity	83.7%	Pred. No. 1.8e-07		
Matches 77	Conservative	0	Mismatches 15	Indels 0
			Gaps	0

Oy 3755 TGGCAGTGAAGCCGAGATCGCACCACTGCACCCCCTGGCTGGCGGCACAGAGTTGACTTTGT 3814
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21403 TTGCAGTGAAGCCGAGATCGCACCACTGCACCTTGCGCAACAGAGTGAAGACTCTGT 2146

Qy 3815 CTCTATTACAAAAAGAAAAGAAATACA 3846
||| | |||| |||| ||| ||| |
Db 21463 CTCAAAACAAAAAAGAAAAAAGAGA 21494

```

RESULT 10
US-09-426-290-1/C
; Sequence 1, Application US/09426250
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglynd Ran Olafsdottir
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FASTSEQ For Windows Version 4.0.
; SEQ ID NO 1

```

```

1 ORGANISM: Homo sapiens
2
3 FEATURE:
4
5 NAME/KEY: CDS
6 LOCATION: (21181) ... (21403)
7
8 NAME/KEY: CDS
9 LOCATION: (95252) ... (95430)
10
11 NAME/KEY: CDS
12 LOCATION: (101753) ... (101966)
13
14 NAME/KEY: CDS
15 LOCATION: (110324) ... (110439)
16
17 NAME/KEY: CDS
18 LOCATION: (124058) ... (124278)
19
20 NAME/KEY: CDS
21 LOCATION: (127009) ... (127130)
22
23 NAME/KEY: CDS
24 LOCATION: (128910) ... (129139)
25
26 OS-09-426-290-1

```

Query Match	1.88; Score 67.8; DB 4; Length 168575
-------------	---------------------------------------

```

Best Local Similarity  /8.00;  Freq. NO. 4.42e-07;
Matches 81;  Conservative 0;  Mismatches 22;  Indels 0;  Gaps 0;

```

3755 TGGCAGTGAAGCGAGATTGGCACCACCTGCACCCCTGCCTGGGGCAGACAGAGTGAAGATTGGT 3814

Db 86451 TTGCAGTAGCGGAGATGCTGCACCTGCTCCAGCCCTGGGCGACAGAGTGAAGTCCGT 86392

Qy 3815 CTCATTATACAAAAGAAAAGAAATACACCTTAAGTCA 3857

Db 86391 CTCAAAAAAGAAAAGAAAAGAAAATTTACCGCTATGCACA 86349

RESULT 11

US-09-078-294-3/c
; Sequence 3, Application US/09078294
; Patent No. 6265211

GENERAL INFORMATION:

APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Salt, Desiree
APPLICANT: Cancilla, Michael R.
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
FILE REFERENCE: Davies Col
CURRENT APPLICATION NUMBER: US/09/078,294
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 80595
TYPE: DNA
ORGANISM: Nucleotide sequence of HC-contig

US-078-294-3

Query Match 1.7%; Score 67.2; DB 4; Length 80595;
Best Local Similarity 72.5%; Pred. No. 4.1e-07;
Matches 87; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 3723 CAGCATATATCCCTAGAAATACACCTAATGGCAGTGGCGAGATGCCACACCTGC 3782

Db 5032 CAGGGAATCACTTGAACCCAGGCGAGAGTTGCAGTGACGAGATCGCGCACTGC 4973

Qy 3783 ACCCTGGCTGGGCGACAGAGTGTCTCTATTACAAAAGAAAAGAAA 3842

Db 4972 ACTCAGGCTGGGTGACAGAGGAGACACTCTGTCAAAAAAAGAAAAGAAA 4913

RESULT 12

US-09-128-155-17/c
; Sequence 17, Application US/09128155
; Patent No. 6117654

GENERAL INFORMATION:

APPLICANT: Pan, Yangu
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 1.7%; Score 67; DB 3; Length 176373;
Best Local Similarity 79.8%; Pred. No. 7e-07;
Matches 79; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 3752 AATTGGACATGAGCGAGATGCACACCTGCTGGGCGACAGAGTGAAGTCTT 3811

Db 30773 AACTTGACGTAGCTGAGATGCGGCACCTGCACCTCCAGCCCTGGGCGACAGAGTGAAGTCTT 30714

Qy 3812 TGCTCTATTTACAAAAGAAAAGAAAAGAAATACACCT 3850

Db 30713 CGTCTCAAAAAAAGAAAAGAAAAGAAAATGATTAACACCT 30675

RESULT 13

US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188

GENERAL INFORMATION:

APPLICANT: Zianphe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
FILE REFERENCE: CL001151CIP
CURRENT APPLICATION NUMBER: US/09/797,906
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 84495
TYPE: DNA
ORGANISM: Human

FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(84495)
OTHER INFORMATION: n = A,T,C or G

US-09-797-906-3

Query Match 1.7%; Score 66.6; DB 4; Length 84495;
Best Local Similarity 84.3%; Pred. No. 6e-07;
Matches 75; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3755 TGGCAGTGGCGGAGATGCGACACCTGCTGGGCGACAGAGTGAAGTCTTGT 3814

Db 20107 TTGTAGTGTGAGATGAGATGCGACACCTGCACCTCGGCGAGAGTGAAGTCTTGT 20166

RESULT 14

US-08-965-048-5/c
; Sequence 5, Application US/08965048
; Patent No. 6323244

GENERAL INFORMATION:

APPLICANT: Freimer, Nelson
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR WHI DIAGNOSIS AND
TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
FILE REFERENCE: 7853-093
CURRENT APPLICATION NUMBER: US/08/965,048
CURRENT FILING DATE: 1997-11-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 45716
TYPE: DNA
ORGANISM: Homo sapiens

US-08-965-048-5

Query Match 1.7%; Score 66.4; DB 4; Length 45716;
Best Local Similarity 71.0%; Pred. No. 4.9e-07;
Matches 88; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 3723 CAGCATATATCTCTAGAAATACACCTAATGGCAGTGGCGAGATGCCACACCTGC 3782

Db 3997 CAGGGAATCGCTTGAACCCAGGCGAGGCTTGCAGTGACGAGATCGCACCACTGC 3938

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:52:35 : Search time 110.711 Seconds
(without alignments)
13416.829 Million cell updates/sec

Title: US-09-919-497-14

Perfect score: 3857
Sequence: 1 ggaagagcttcacgtgagtc.....agaataacacctaagctca 3857

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 341543 seqs, 192557720 residues

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US06_PUB_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3857	100.0	3857	10	US-09-919-497-14
2	2857.4	74.1	2900	10	US-09-925-300-670
3	1829.4	47.4	35641	10	US-09-962-436-306
4	1829.4	47.4	35641	10	US-09-880-107-2225
5	256.2	6.6	601	10	US-09-867-701-10472
6	154.6	4.0	35641	10	US-09-919-497-14
7	154.6	4.0	35641	10	US-09-962-436-306
8	154.6	4.0	35641	10	US-09-880-107-2225
9	70.2	1.8	9566	10	US-09-764-864-1768
10	70.2	1.8	9566	10	US-09-764-864-1768
11	69.2	1.8	3705	10	US-09-764-877-3486
12	69.2	1.8	145831	10	US-09-969-708-79
13	69.2	1.8	145831	10	US-09-954-456-2116
14	69.2	1.8	24218	9	US-09-860-670-263
15	68	1.8	323	10	US-09-764-878-410
16	68	1.8	323	10	US-09-764-878-410
17	68	1.8	5299	10	US-09-764-878-409
18	68	1.8	8210	10	US-09-764-878-406
19	67.6	1.8	1400	10	US-09-263-959-295

C	20	67.6	1.8	9474	10	US-09-764-878-372	Sequence 372, App
C	21	67.6	1.8	684973	10	US-09-263-959-1	Sequence 1, Appl
C	22	67.4	1.7	279	10	US-09-764-887-507	Sequence 507, App
C	23	67.4	1.7	1668	10	US-09-993-811-13	Sequence 13, Appl
C	24	67.4	1.7	4962	10	US-09-764-869-1902	Sequence 1902, App
C	25	67.4	1.7	4962	10	US-09-764-869-1903	Sequence 1903, App
C	26	67.4	1.7	9192	10	US-09-764-847-1246	Sequence 1246, App
C	27	67.4	1.7	9192	10	US-09-764-847-1246	Sequence 1246, App
C	28	67.4	1.7	9745	10	US-09-764-869-2258	Sequence 2258, App
C	29	67.4	1.7	12149	10	US-09-764-869-2258	Sequence 2258, App
C	30	67.4	1.7	13182	10	US-09-817-1994-3	Sequence 3, Appl
C	31	67.4	1.7	32203	10	US-09-764-869-1849	Sequence 1849, App
C	32	67.4	1.7	368004	10	US-09-949-654-3	Sequence 3, Appl
C	33	67.2	1.7	323	10	US-09-867-701-9630	Sequence 9630, App
C	34	67.2	1.7	23603	9	US-09-860-670-264	Sequence 264, App
C	35	67.2	1.7	23613	9	US-09-860-670-258	Sequence 258, App
C	36	67.2	1.7	24699	10	US-09-764-877-2419	Sequence 2419, App
C	37	67.2	1.7	167343	10	US-09-962-436-281	Sequence 281, App
C	38	67.2	1.7	167343	10	US-09-964-824A-273	Sequence 273, App
C	39	67.2	1.7	21833	10	US-09-764-877-2275	Sequence 2275, App
C	40	67	1.7	31994	9	US-09-764-904-71	Sequence 71, Appl
C	41	67	1.7	31994	9	US-09-764-860-599	Sequence 599, App
C	42	67	1.7	176373	9	US-10-095-407-17	Sequence 17, Appl
C	43	66.8	1.7	329	10	US-09-764-877-3990	Sequence 3990, App
C	44	66.6	1.7	565	10	US-09-925-301-611	Sequence 611, App
C	45	66.6	1.7	2364	10	US-09-764-887-513	Sequence 513, App

ALIGNMENTS

RESULT 1
US-09-919-497-14
Sequence 14, Application US/09919497

Patent No. US2002010662A1
GENERAL INFORMATION:
APPLICANT: Muller, George L.
TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
FILE REFERENCE: B0801/7225
CURRENT APPLICATION NUMBER: US/09/919, 497
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/221, 735
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 100
SOFTWARE: Patent version 3.0
SEQ ID NO 14
LENGTH: 3857
TYPE: DNA
ORGANISM: Homo sapiens
US-09-919-497-14

Query Match 100.0%; Score 3857; DB 10; Length 3857;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAAGAGCTTCATGAGTCAGAGGACCGGCTGACAGGATTCATGGCGCGCTGCG	60
DB	1	GGAAGAGCTTCATGAGTCAGAGGACCGGCTGACAGGATTCATGGCGCGCTGCG	60
QY	61	GAGTACGAGCGACACCTGCTGAGTCTGACAGTCTGACAGTCTGAGTCTGAGTCTG	120
DB	61	GAGTACGAGCGACACCTGCTGAGTCTGACAGTCTGACAGTCTGAGTCTGAGTCTG	120
QY	121	CGCGGCTGCGGCGGACCGGCTGCTTACACATCTTCTTACAGTCTGACAGTCTG	180
DB	121	CGCGGCTGCGGCGGACCGGCTGCTTACACATCTTCTTACAGTCTGACAGTCTG	180
QY	181	TGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240
DB	181	TGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240
QY	241	AGATAGAGGAGTGAACACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300
DB	241	AGATAGAGGAGTGAACACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300

Db 241 AAGATGAAGAGTGTGAACACCTCCCTCCGCTGTAAACAATGAATCAACAAGCAGGT 300
QY 301 CGCTATGAGAGTTACACAGAGGTGCTGTTATTTTGGCAGCAATGAGTACTTGGT 360
Db 301 CGCTATGAGAGTTACACAGAGGTGCTGTTATTTTGGCAGCAAGTACTTGGT 360
QY 361 GACTCTTGACATGATGAATACCTTCAGATTTTAACTGCGACTTGGTGTATAGACC 420
Db 361 GACTCTTGACATGATGAATACCTTCAGATTTTAACTGCGACTTGGTGTATAGACC 420
QY 421 CACAGATATGACGCTGTTGTCAAGAGCAATTCCTGGCGCTTTGGCCAGAAAAAC 480
Db 421 CACAGATATGACGCTGTTGTCAAGAGCAATTCCTGGCGCTTTGGCCAGAAAAAC 480
QY 481 AAGCGTGTATTTTAAAGCTTTCACAGCAATGCTTGGCTTTGATACGTGTTTTGT 540
Db 481 AAGCGTGTATTTTAAAGCTTTCACAGCAATGCTTGGCTTTGATACGTGTTTTGT 540
QY 541 CATGTGGAAGAGTATGAGAAATCTTTTGTGAGAAACTGTATCTGTGGCCAGAAAAAC 600
Db 541 CATGTGGAAGAGTATGAGAAATCTTTTGTGAGAAACTGTATCTGTGGCCAGAAAAAC 600
QY 601 CATGTGAGCAATGATATTTTGAAGCAACCAACCTGAAGTGTAGAAATCCATGTT 660
Db 601 CATGTGAGCAATGATATTTTGAAGCAACCAACCTGAAGTGTAGAAATCCATGTT 660
QY 661 TCTATGACACCTTACATGCTTGTCTATACAGACTGTATCTGACATTTTAAATGCAATG 720
Db 661 TCTATGACACCTTACATGCTTGTCTATACAGACTGTATCTGACATTTTAAATGCAATG 720
QY 721 CTAAAGGAACTAAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 780
Db 721 CTAAAGGAACTAAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 780
QY 781 GCTATTGGAAGAACTTTTGGCAAGACATCCGCTTATCTGATCTTGGCCAGCAG 840
Db 781 GCTATTGGAAGAACTTTTGGCAAGACATCCGCTTATCTGATCTTGGCCAGCAG 840
QY 841 CTGTGAGCCAAAGATTAATCTTATGATTCAGATTTGAAGATTTAGCAACTTGTGTCAG 900
Db 841 CTGTGAGCCAAAGATTAATCTTATGATTCAGATTTGAAGATTTAGCAACTTGTGTCAG 900
QY 901 TATCTGCTAGATGATGATGTCATGATTTCTTAATCTCTGGAATCTGAGAGCAAG 960
Db 901 TATCTGCTAGATGATGATGTCATGATTTCTTAATCTCTGGAATCTGAGAGCAAG 960
QY 961 GAAAAAGCTTTTGGTGAAGATTCAGATTCAGATTCCTTCTGACTCCAGCACTGATGTT 1020
Db 961 GAAAAAGCTTTTGGTGAAGATTCAGATTCAGATTCCTTCTGACTCCAGCACTGATGTT 1020
QY 1021 ATTAATGCTGAGCAAGGCTTATATCTTCCAGATGCCAAATGAGTAAAGAAAAA 1080
Db 1021 ATTAATGCTGAGCAAGGCTTATATCTTCCAGATGCCAAATGAGTAAAGAAAAA 1080
QY 1081 ATATCTGAAAAAATGAAAAATTAAGAAGGAGAAAAAAGAAAAAGAACTGCTCTAGAA 1140
Db 1081 ATATCTGAAAAAATGAAAAATTAAGAAGGAGAAAAAAGAAAAAGAACTGCTCTAGAA 1140
QY 1141 AGCAACCCAAAGTGGAGGAGCTGATGAATTTAAAGAAATTAAGGAGAAATTAAG 1200
Db 1141 AGCAACCCAAAGTGGAGGAGCTGATGAATTTAAAGAAATTAAGGAGAAATTAAG 1200
QY 1201 GAGAGTGAAGCTTGTGTGTCAGGTCAAGTACTGATTTGTGTCAGAGTGTAGCCGAAC 1260
Db 1201 GAGAGTGAAGCTTGTGTGTCAGGTCAAGTACTGATTTGTGTCAGAGTGTAGCCGAAC 1260
QY 1261 TGTTCCTGAGTGAAGACTATATCACTTTTGGAGCGAGGCTTTTATTTAGAGCTCTAC 1320
Db 1261 TGTTCCTGAGTGAAGACTATATCACTTTTGGAGCGAGGCTTTTATTTAGAGCTCTAC 1320
QY 1321 AGCAAAACCTTTGAGAGAGTATGCAAAAGCTGAAGAGTGTGATGAAATTTAGCAAGGAA 1380
Db 1321 AGCAAAACCTTTGAGAGAGTATGCAAAAGCTGAAGAGTGTGATGAAATTTAGCAAGGAA 1380

QY 1381 GACAGTTCAAGAGATTTAGAAATCTACAAAAAGACTTAAGACCCCAAAAAAGAA 1440
Db 1381 GACAGTTCAAGAGATTTAGAAATCTACAAAAAGACTTAAGACCCCAAAAAAGAA 1440
QY 1441 CGGGCTTCTACCAAGAAAGAACCCCTCAAAAAAGAAAAAGAACTTGAACCTTAACTCA 1500
Db 1441 CGGGCTTCTACCAAGAAAGAACCCCTCAAAAAAGAAAAAGAACTTGAACCTTAACTCA 1500
QY 1501 ATGCTAGAAAAACCTTAAGTGAAGAGAGAGATGTGAGTGAAGATATGCTCA 1560
Db 1501 ATGCTAGAAAAACCTTAAGTGAAGAGAGAGATGTGAGTGAAGATATGCTCA 1560
QY 1561 GAAATTAAGCAATGAGCCCAAGAAAGCTCCGGAAGAAATTAAGCAATGAAATTTGATGA 1620
Db 1561 GAAATTAAGCAATGAGCCCAAGAAAGCTCCGGAAGAAATTAAGCAATGAAATTTGATGA 1620
QY 1621 AATTTGTCATGCGATGCTGCTTTGGAATCTGAAAGAACCCCTCAATTCATCCAG 1680
Db 1621 AATTTGTCATGCGATGCTGCTTTGGAATCTGAAAGAACCCCTCAATTCATCCAG 1680
QY 1681 CTTCGTGGTTGACGAGCCCTATGCTGTGACAAAGGCTACATGAGTGAAGCCAGA 1740
Db 1681 CTTCGTGGTTGACGAGCCCTATGCTGTGACAAAGGCTACATGAGTGAAGCCAGA 1740
QY 1741 TACGTGTTCTTTATGACGACAGAGTAACTTGTTCGACACTTGAATTTTACAGGGCG 1800
Db 1741 TACGTGTTCTTTATGACGACAGAGTAACTTGTTCGACACTTGAATTTTACAGGGCG 1800
QY 1801 AGTAGGCTTGGGAACCTGTGAGAGTTTACTTCTTATATAGGAGTCAACTGAGAA 1860
Db 1801 AGTAGGCTTGGGAACCTGTGAGAGTTTACTTCTTATATAGGAGTCAACTGAGAA 1860
QY 1861 CAAGGCTATCTCACTGCTTTTGGGGAAGAAAGAAAGCTTTTGAAGAACTATAGGGA 1920
Db 1861 CAAGGCTATCTCACTGCTTTTGGGGAAGAAAGAAAGCTTTTGAAGAACTATAGGGA 1920
QY 1921 AAAGCAAGATGTTTCTCCCTGAAGAAAGAAAGGAGATGAACAACCTTACACTTA 1980
Db 1921 AAAGCAAGATGTTTCTCCCTGAAGAAAGAAAGGAGATGAACAACCTTACACTTA 1980
QY 1981 GTTAAGAGGACAGCATCTGAGATGTTTCCACTGACATCGGAACCCGCTGGCCAGAA 2040
Db 1981 GTTAAGAGGACAGCATCTGAGATGTTTCCACTGACATCGGAACCCGCTGGCCAGAA 2040
QY 2041 CAGAAATGTACACAGCAAAAGCATAGTTGTGATATGCTGAATTTGAAAGTGAAGCTTCA 2100
Db 2041 CAGAAATGTACACAGCAAAAGCATAGTTGTGATATGCTGAATTTGAAAGTGAAGCTTCA 2100
QY 2101 TCTCTGATCATGCTGGGGCATTTGACATTTGAACCCGTACTTTATAGTTGAGATTTAC 2160
Db 2101 TCTCTGATCATGCTGGGGCATTTGACATTTGAACCCGTACTTTATAGTTGAGATTTAC 2160
QY 2161 ATCCCTACATCGAAGATGTCGAGGCAAGAGATGATGATGATTTATTCGGCTTTTA 2220
Db 2161 ATCCCTACATCGAAGATGTCGAGGCAAGAGATGATGATGATTTATTCGGCTTTTA 2220
QY 2221 AATTAAGGCGGCTCTACAGCCAGTGCATCTCCATCTCCCGCTACTTCAAGCGTCCCGT 2280
Db 2221 AATTAAGGCGGCTCTACAGCCAGTGCATCTCCATCTCCCGCTACTTCAAGCGTCCCGT 2280
QY 2281 CTTCGATTTGAGTTGACCTTACCAAGCTTTCTCTCTCACTTCCGAGGTGCTTTGTT 2340
Db 2281 CTTCGATTTGAGTTGACCTTACCAAGCTTTCTCTCTCTCACTTCCGAGGTGCTTTGTT 2340
QY 2341 CAGGATATCTCCAGCAATGATATGTTTCCAAATCACTCTTCTTACACTTCACTTCCCG 2400
Db 2341 CAGGATATCTCCAGCAATGATATGTTTCCAAATCACTCTTCTTACACTTCACTTCCCG 2400
QY 2401 AGACTACGATTTCTGTGGCCCTCTCTCATGCAAGGCGAGTGTGTTGAGAGCTG 2460
Db 2401 AGACTACGATTTCTGTGGCCCTCTCTCATGCAAGGCGAGTGTGTTGAGAGCTG 2460

QY 2461 AAACAAAGCAGCCAGCAGCTGATGCGGCGACAGCAGCTGGCCATTACAGCAAGATTCGTGA 2520
 DB 2461 AAACAAAGCAGCCAGCAGCTGATGCGGCGACAGCAGCTGGCCATTACAGCAAGATTCGTGA 2520
 QY 2521 ACCCTCCGAGCTCAGAGAGATATATCTGCTCCCAAGACTTCTGTATAAATGCCA 2580
 DB 2521 ACCCTCCGAGCTCAGAGAGATATATCTGCTCCCAAGACTTCTGTATAAATGCCA 2580
 QY 2581 GGGGTGAATGCCAAAACCTGCGCTCTTGATGCACACAGCTTGAAGACATCGCAGATTA 2640
 DB 2581 GGGGTGAATGCCAAAACCTGCGCTCTTGATGCACACAGCTTGAAGACATCGCAGATTA 2640
 QY 2641 GCAGCCCTGTCAAGAGAGCTCAGAGATTTGGGGAAATGCTGCAATGCCAAACAG 2700
 DB 2641 GCAGCCCTGTCAAGAGAGCTCAGAGATTTGGGGAAATGCTGCAATGCCAAACAG 2700
 QY 2701 CTTTATGATTTTATTCACACCTCTTTTGCAGAGTCGTATCAAAAAGAAAGGAAAAAG 2760
 DB 2701 CTTTATGATTTTATTCACACCTCTTTTGCAGAGTCGTATCAAAAAGGAAAAAG 2760
 QY 2761 TGAACAGTATGCTGTTTCTTATCCCATGCTGACTTTTCAGGGGCTCTTGCCAGA 2820
 DB 2761 TGAACAGTATGCTGTTTCTTATCCCATGCTGACTTTTCAGGGGCTCTTGCCAGA 2820
 QY 2821 CATCATAGCTCATTTATTAATTAATGTTGCTATTTTCATTTCTTTCCAAATGCTTTAATG 2880
 DB 2821 CATCATAGCTCATTTATTAATTAATGTTGCTATTTTCATTTCTTTCCAAATGCTTTAATG 2880
 QY 2881 ATTGTACGCTGACCAAGAGCCAGATTCCTCTGAACTCTGCAGTTAGGCATCTTG 2940
 DB 2881 ATTGTACGCTGACCAAGAGCCAGATTCCTCTGAACTCTGCAGTTAGGCATCTTG 2940
 QY 2941 AACTTGCCTGCTGCTGCTCTTTTCTCTCCCTGACGCTGATGCGGGGCTTGACATGTT 3000
 DB 2941 AACTTGCCTGCTGCTGCTCTTTTCTCTCCCTGACGCTGATGCGGGGCTTGACATGTT 3000
 QY 3001 CTTTTTAATGAGGTTTGTACAGATCAGGTAAGTTCCTACAAGTATTCAGAAAGTAG 3060
 DB 3001 CTTTTTAATGAGGTTTGTACAGATCAGGTAAGTTCCTACAAGTATTCAGAAAGTAG 3060
 QY 3061 AAACCTTACCTGATCTCTAAGATCTCATTTGAAAAGAAATGCTAACCTGCGCATGA 3120
 DB 3061 AAACCTTACCTGATCTCTAAGATCTCATTTGAAAAGAAATGCTAACCTGCGCATGA 3120
 QY 3121 CGGTGAGGAGGAGAAAGAGGAGCAGACAGAAAGCTACATTTTAAAGCTCTGTGA 3180
 DB 3121 CGGTGAGGAGGAGAAAGAGGAGCAGACAGAAAGCTACATTTTAAAGCTCTGTGA 3180
 QY 3181 TCTAGTGCACATTAATAACACTCTTAATGCACTTATACCATGCTCTGCTGCCA 3240
 DB 3181 TCTAGTGCACATTAATAACACTCTTAATGCACTTATACCATGCTCTGCTGCCA 3240
 QY 3241 AATCTGCTTGTCTGTCTGCTGCTGAGCCTTGAAGTGTGTTGTGTAAGAAATCA 3300
 DB 3241 AATCTGCTTGTCTGTCTGCTGCTGAGCCTTGAAGTGTGTTGTGTAAGAAATCA 3300
 QY 3301 TGTCTGACCCCTTGTCTACAAAGAGCCTTGTGAAACCTGAGAAGAAACATCTTTG 3360
 DB 3301 TGTCTGACCCCTTGTCTACAAAGAGCCTTGTGAAACCTGAGAAGAAACATCTTTG 3360
 QY 3361 CCATTCCTGACCAAGTGTCTCTACACATTTTCTCAGCTCATCTCTGCTGCTGTC 3420
 DB 3361 CCATTCCTGACCAAGTGTCTCTACACATTTTCTCAGCTCATCTCTGCTGCTGTC 3420
 QY 3421 TCTAAGGAAATTTTCTAGAGCCTTCTCTACTAATTCAGACAGCTCTCTCAAAACG 3480
 DB 3421 TCTAAGGAAATTTTCTAGAGCCTTCTCTACTAATTCAGACAGCTCTCTCAAAACG 3480
 QY 3481 GTTGACTAGCTCTTAATGACCCCTAACATATGTAGCATTTACTATTAATTTCTGTTCA 3540
 DB 3481 GTTGACTAGCTCTTAATGACCCCTAACATATGTAGCATTTACTATTAATTTCTGTTCA 3540
 QY 3541 AATTAGTATTTTAAAGCAAAATGAATTACCTGTTGCAAAAAGTTAATGATGAAGAGCT 3600

DB 3541 AATTAGTATTTTAAAGCAAAATGAATTACCTGTTTGCAAAAGTTAATGATGAAGAGCT 3600
 QY 3601 CTTAGATTCATATTTTGGACATATTCAGTCTCTATATATACAGATCCCTAAGTCA 3660
 DB 3601 CTTAGATTCATATTTTGGACATATTCAGTCTCTATATATACAGATCCCTAAGTCA 3660
 QY 3661 GCTGGCTAGTACAGAGTTTTCAGACTTCTCTGTTCTCAGCTCTTATCTTAAGAC 3720
 DB 3661 GCTGGCTAGTACAGAGTTTTCAGACTTCTCTGTTCTCAGCTCTTATCTTAAGAC 3720
 QY 3721 ACCAGCATATATCTCTAGAAATACAACTAATTTGGCAGTGAAGCCAGATCCAGCACT 3780
 DB 3721 ACCAGCATATATCTCTAGAAATACAACTAATTTGGCAGTGAAGCCAGATCCAGCACT 3780
 QY 3781 GCACCCCTGCTGGGCGACAGAGTGAAGCTTTGCTCTATTACAAAAAGAAAGAA 3840
 DB 3781 GCACCCCTGCTGGGCGACAGAGTGAAGCTTTGCTCTATTACAAAAAGAAAGAA 3840
 QY 3841 AATACAACTAAGCTCA 3857
 DB 3841 AATACAACTAAGCTCA 3857

RESULT 2
 US-09-925-300-670
 ; Sequence 670, Application US/09925300
 ; Patent No. US20020151681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig Rosen,
 ; APPLICANT: Steve Ruben
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA101
 ; CURRENT APPLICATION NUMBER: US/09/925,300
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; NUMBER OF SEQ ID NOS: 1890
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 670
 ; LENGTH: 2900
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (2418)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-925-300-670

Query Match 74.1%; Score 2857.4; DB 10; Length 2900;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2869; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
 QY 25 CAGCGGCTCGAGAGGATTCATGCGGCGCTGAGTACAGCAGAGTGGGCTG 84
 DB 15 CCGCGGCTCGAGAGGATTCATGCGGCGCTGAGTACAGCAGAGTGGGCTG 73
 QY 85 GAATGCTCGACACTGAGCGGCTAGTAGTGTGCGGCGGCGCTGCGCGGAGCGGCTC 144
 DB 74 GAATGCTCGACACTGAGCGGCTAGTAGTGTGCGGCGGCGCTGCGCGGAGCGGCTC 133
 QY 145 CTTTACCACTTTCTCAGTGCACGTGCAAGCCAGCTGCTGCTGCTGCTCAACAG 204
 DB 134 CTTTACCACTTTCTCAGTGCACGTGCAAGCCAGCTGCTGCTGCTGCTCAACAG 193
 QY 205 CAGCGGCGGAGAGAGATTTTATCAATCAGCTGAAGATGAAGAGACTTAACACCTC 264
 DB 194 CAGCGGCGGAGAGAGATTTTATCAATCAGCTGAAGATGAAGAGACTTAACACCTC 253
 QY 265 CCTGCGCTGTAACAAATGAATCAACAAAGCAAGTGGCTATGAAGTTTACACAGAGCT 324

Db 254 CCTGCGCCGTGAACAAATGAAATCACAGCAACAGTCGTATGAAATTTACACCAAGT 313
QY 325 GGTGTTATATTGGCAAAATAGATACTTGTGGTTGACTTCTTGATGATAGAAATACCT 384
Db 314 GGTGTTATATTGGCAAAATAGATACTTGTGGTTGACTTCTTGATGATAGAAATACCT 373
QY 385 TCAGATTTAAATTACGCAATCTTGCTGATAGAGCCACAGAAATATGAGATCTTGCA 444
Db 374 TCAGATTTAAATTACGCAATCTTGCTGATAGAGCCACAGAAATATGAGATCTTGCA 433
QY 445 GAAGCATTCATCTTGCCCTCTTTGCCAGAAAAACAACAGTGGTTTATTAAGCTTTC 504
Db 434 GAAGCATTCATCTTGCCCTCTTTGCCAGAAAAACAACAGTGGTTTATTAAGCTTTC 493
QY 505 ACAGACAATGCTGTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 564
Db 494 ACAGACAATGCTGTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 553
QY 565 CTTTGTGAGGAACATGATATCTGTGGCAAGTTCCATGATAGCAATCACTATTTT 624
Db 554 CTTTGTGAGGAACATGATATCTGTGGCAAGTTCCATGATAGCAATCACTATTTT 613
QY 625 GAACAGACAACCTGAAAGTTGAGAAATCCATGTTCTATGACACCTTACCAGTCTGCT 684
Db 614 GAACAGACAACCTGAAAGTTGAGAAATCCATGTTCTATGACACCTTACCAGTCTGCT 673
QY 685 ATACAGCTGCTATACGCAATTTTAAATGATGATGATGATGATGATGATGATGATGAT 744
Db 674 ATACAGCTGCTATACGCAATTTTAAATGATGATGATGATGATGATGATGATGATGAT 733
QY 745 CCATCGCTTGAAGTGAAGATTTATCTTGAAGAAATGCTATTGGAAAAACCTTTGACAG 804
Db 734 CCATCGCTTGAAGTGAAGATTTATCTTGAAGAAATGCTATTGGAAAAACCTTTGACAG 793
QY 805 ACAATCCGCAATATCTGATCCTTTGTGCACACAGCTTGGAGCCAAAGACTAAATCTTA 864
Db 794 ACAATCCGCAATATCTGATCCTTTGTGCACACAGCTTGGAGCCAAAGACTAAATCTTA 853
QY 865 GTTCAGATTTGAAGATATTAAGCAATTTGCTGAGATCTCTCTGATGATGATGATGATG 924
Db 854 GTTCAGATTTGAAGATATTAAGCAATTTGCTGAGATCTCTCTGATGATGATGATGATG 913
QY 925 ACATCTTATATCTTCTGATCCTGAGACCAACGAAAAAGCTTTTGGTCAGAAATTA 984
Db 914 ACATCTTATATCTTCTGATCCTGAGACCAACGAAAAAGCTTTTGGTCAGAAATTA 973
QY 985 GGTGGCTGTTTCTGACTCCAGACCTCGATGTTTAAATGCTCGAGCAAGGGTTTAT 1044
Db 974 GGTGGCTGTTTCTGACTCCAGACCTCGATGTTTAAATGCTCGAGCAAGGGTTTAT 1033
QY 1045 CATCTTCAGATGCAAAATGAGTAAAAAAGAAATATCTGAAAAATGGAATTTAA 1104
Db 1034 CATCTTCAGATGCAAAATGAGTAAAAAAGAAATATCTGAAAAATGGAATTTAA 1093
QY 1105 GAAGGGGAAGAAACAAAAAGAACTGCTCTAGAAAAAGCAACCAAGCTGGAGGCACTG 1164
Db 1094 GAAGGGGAAGAAACAAAAAGAACTGCTCTAGAAAAAGCAACCAAGCTGGAGGCACTG 1153
QY 1165 ACTGAATATTTAAAGAAATGAGGCAAGAAATTAAGAGAGTGAAGCTTTGGTGCTCA 1224
Db 1154 ACTGAATATTTAAAGAAATGAGGCAAGAAATTAAGAGAGTGAAGCTTTGGTGCTCA 1213
QY 1225 GGTCAATGATGATTTGCAAGTATGACGAAATGTTCCAGCTGAGAGATATATC 1284
Db 1214 GGTCAATGATGATTTGCAAGTATGACGAAATGTTCCAGCTGAGAGATATATC 1273
QY 1285 ACTCTGAGAGGAGGCTTTCTATTGAGGCTCTACAGAAAACTTTGAGAGATAGC 1344
Db 1274 ACTCTGAGAGGAGGCTTTCTATTGAGGCTCTACAGAAAACTTTGAGAGATAGC 1333
QY 1345 AAAGCTGAGAGAGTCTGATGAATTTAGAGAGAGAGACAGTTCAAAAGAGATTTAGAAA 1404
Db 1334 AAAGCTGAGAGAGTCTGATGAATTTAGAGAGAGAGACAGTTCAAAAGAGATTTAGAAA 1393
QY 1405 TCTCAGAAAAGACCTTAAGAACCCCAAAAGCAAGAGGCTTCTATCAAAAGAAAGAAC 1464
Db 1394 TCTCAGAAAAGACCTTAAGAACCCCAAAAGCAAGAGGCTTCTATCAAAAGAAAGAAC 1453
QY 1465 CTCAAAAAGAAAAAGGAGGATGACCTTAATCTCAAAATGATAGAAACCTGAAAGACTG 1524
Db 1454 CTCAAAAAGAAAAAGGAGGATGACCTTAATCTCAAAATGATAGAAACCTGAAAGACTG 1513
QY 1525 GAAGAGAGAGAGATGTCGAGAGAGATATCTGAGAAATTAAGCAATAGCCAGAAAGC 1584
Db 1514 GAAGAGAGAGAGATGTCGAGAGAGATATCTGAGAAATTAAGCAATAGCCAGAAAGC 1573
QY 1585 TGCCCGAAGAAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1644
Db 1574 TGCCCGAAGAAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1633
QY 1645 GGAATCCTGAAAAGAACCCCTCATATCATCCATCCGCTTGGGTTGACCGACCCCTAT 1704
Db 1634 GGAATCCTGAAAAGAACCCCTCATATCATCCATCCGCTTGGGTTGACCGACCCCTAT 1693
QY 1705 GCTTGACAAAGGATACATGAAATGAGGCAAGATACGTCGTTCTTTATGACGACAG 1764
Db 1694 GCTTGACAAAGGATACATGAAATGAGGCAAGATACGTCGTTCTTTATGACGACAG 1753
QY 1765 CTACCTTTGTGGGAGCTTGAATTTACAGGGGAGTATGAGGCTTGGGAAACCTGTAGG 1824
Db 1754 CTACCTTTGTGGGAGCTTGAATTTACAGGGGAGTATGAGGCTTGGGAAACCTGTAGG 1813
QY 1825 GTTACTTTCTTATATACGAGGTTTCACTGAGAGAACAGCTATCTCACTGCTTTGCGG 1884
Db 1814 GTTACTTTCTTATATACGAGGTTTCACTGAGAGAACAGCTATCTCACTGCTTTGCGG 1873
QY 1885 AAAGAAAAAGAGCTTTTGAAGAACTATAGGAAAAAGCAAGCATGTTTCTCTGAA 1944
Db 1874 AAAGAAAAAGAGCTTTTGAAGAACTATAGGAAAAAGCAAGCATGTTTCTCTGAA 1933
QY 1945 GAAAGAGAGAGAGATGAAACAACTTAAGCTTAAGAGGCAACAGATGCGAGAT 2004
Db 1934 GAAAGAGAGAGAGATGAAACAACTTAAGCTTAAGAGGCAACAGATGCGAGAT 1993
QY 2005 GTTTCACAGTACATCGGAAAGCGGTGGCCAGAGAACAGATGCTATACAGCAAGCAT 2064
Db 1994 GTTTCACAGTACATCGGAAAGCGGTGGCCAGAGAACAGATGCTATACAGCAAGCAT 2053
QY 2065 GTTGTGATATGCTGATTTTCAAGTGAAGCTTCAATCTCTGATCCTCTCGGGGAT 2124
Db 2054 GTTGTGATATGCTGATTTTCAAGTGAAGCTTCAATCTCTGATCCTCTCGGGGAT 2113
QY 2125 GACATTAACCCGATGATTAAGAGGTTGAGATTAATCCGACCTCAGAAATGTCGCTG 2184
Db 2114 GACATTAACCCGATGATTAAGAGGTTGAGATTAATCCGACCTCAGAAATGTCGCTG 2173
QY 2185 GAGCGAAGATATCACTGATTAATCGGCTTTTAAATTAAGCGCCCTCTACAGCAG 2244
Db 2174 GAGCGAAGATATCACTGATTAATCGGCTTTTAAATTAAGCGCCCTCTACAGCAG 2233
QY 2245 TGCATCTCAGATGTCGCCCTACTACAGCTTCGCTGCTGATGATGATGATGATGATGAT 2304
Db 2234 TGCATCTCAGATGTCGCCCTACTACAGCTTCGCTGCTGATGATGATGATGATGATGAT 2293
QY 2305 AAGCCTTCTCTCACTTCCCGAGGCTTGTATAGAGATCTCAGCAATGACAT 2364
Db 2294 AAGCCTTCTCTCACTTCCCGAGGCTTGTATAGAGATCTCAGCAATGACAT 2353
QY 2365 AGTTCCAAACTCACTTCTTACACTTCACTTCCCGAGAGATCTCTGTGTCGCC 2424
Db 2354 AGTTCCAAACTCACTTCTTACACTTCACTTCCCGAGAGATCTCTGTGTCGCC 2413
QY 2425 TCTCTCATGCAACGGGAGGATGTTTGAAGAGCTGAAACAAAGCAAGCAGACCTGAT 2484
Db 2414 TCTCTCATGCAACGGGAGGATGTTTGAAGAGCTGAAACAAAGCAAGCAGACCTGAT 2473

2485 GCGGAGACAGACATGGCCATTACAGACAGATTCTGAAACCCCTTCCCGAGTCAGAGATAT 2544
|||||
Db 2474 GCGGAGACAGACATGGCCATTACAGACAGATTCTGAAACCCCTTCCCGAGTCAGAGATAT 2533
2545 AATCTGTCTCCCAAGACTTCTGTTAAATGCCAGGGGTGAATGCCAAAACCTGCCGC 2604
|||||
Db 2534 AATCTGTCTCCCAAGACTTCTGTTAAATGCCAGGGGTGAATGCCAAAACCTGCCGC 2593
2605 TCCCTGAGACACACAGTAAAGACATCGCAATATGACAGCCCTGTACAGACAGCTC 2664
|||||
Db 2594 TCCCTGAGACACACAGTAAAGACATCGCAATATGACAGCCCTGTACAGACAGCTC 2653
2665 ACAGATATTCGGGAGATGTCGCAATGCCAAACAGCTTTATGATTTATTCACACCTC 2724
|||||
Db 2654 ACAGATATTCGGGAGATGTCGCAATGCCAAACAGCTTTATGATTTATTCACACCTC 2713
2725 TTTGCGAAGTCGTATCAAAAGGAAAGGAAAGGAAAGTGAACAGTATGGCTGTTCTTA 2784
|||||
Db 2714 TTTGCGAAGTCGTATCAAAAGGAAAGGAAAGGAAAGTGAACAGTATGGCTGTTCTTA 2773
2785 TCCCATGCTGTACTTTTCAGGGGCTCTTGCCAGACATCATAGTATTTATTTATTT 2844
|||||
Db 2774 TCCCATGCTGTACTTTTCAGGGGCTCTTGCCAGACATCATAGTATTTATTTATTT 2833
2845 GGTTCGTAATTCATTTCTTTCGAATGCTCTTAATGATTTGACGTGACGACGAGCGAG 2904
|||||
Db 2834 GGTTCGTAATTCATTTCTTTCGAATGCTCTTAATGATTTGACGTGACGAGCTTCAG 2893

RESULT 3

US-09-962-436-306
; Sequence 306, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 306
; LENGTH: 35641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-306

Query Match 47.4%; Score 1829.4; DB 10; Length 35641;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

2011 ACTGACACTCGGAAAGCCGCTGCGCAGAGAGATGTTAGACAGCAAGCATAGTTG 2070
|||||
Db 29522 ACTTACTTTTCTCTGTAGTGGCCAGAGACAGATGTTAGACAGCAAGCATAGTTG 29581
2071 GATATGCTGATTTTCAAGTGAAGTCTTCATCTCTGATTCATCGTGGGGCATTTGACATT 2130
|||||
Db 29582 GATATGCTGATTTTCAAGTGAAGTCTTCATCTCTGATTCATCGTGGGGCATTTGACATT 29641
2131 GAACCCGATCTTAGAGTGGAGATTACATCTCCTACAGAAATGGCTGAGAGGC 2190
|||||
Db 29642 GAACCCGATCTTAGAGTGGAGATTACATCTCCTACAGAAATGGCTGAGAGGC 29701
2191 AAGATATCACTGATTTTAAATCGGCTCTTTAAATAACGGCCGCTCTACAGCAGTGCATC 2250
|||||
Db 29702 AAGATATCACTGATTTTAAATCGGCTCTTTAAATAACGGCCGCTCTACAGCAGTGCATC 29761
2251 TCCATGTCCCGCTACTACAGAGCTCCGCTGCTTCTGATTTGAGTTGACCCCTAGCAAGCT 2310

29762 TCCATGTCCCGCTACTACAGAGCTCCGCTGCTTCTGATTTGACCTTACGCAAGCT 29821
|||||
2311 TTCTCTCTCATTCCCGAGGTCCTGTTTTCAGGAGATCTCCAGCATGACATTAGTTCC 2370
|||||
Db 29822 TTCTCTCTCATTCCCGAGGTCCTGTTTTCAGGAGATCTCCAGCATGACATTAGTTCC 29881
2371 AAACATCCTCTTCTTACACTTACTTCCCGAGACTGAGATTTCTGTGGCCCTCTCCT 2430
|||||
Db 29882 AAACATCCTCTTCTTACACTTACTTCCCGAGACTGAGATTTCTGTGGCCCTCTCCT 29941
2431 CATGCAAGCGGAGGTTGTTGAGAGCTGAAACAAGAACCCACAGCTGATGAGGCG 2490
|||||
Db 29942 CATGCAAGCGGAGGTTGTTGAGAGCTGAAACAAGAACCCACAGCTGATGAGGCG 30001
2491 ACAGCACTGGCCATTACAGACAGATTCTGAACCCCTTCCGAGTCAGAGAAATATCT 2550
|||||
Db 30002 ACAGCACTGGCCATTACAGACAGATTCTGAACCCCTTCCGAGTCAGAGAAATATCT 30061
2551 GGTCCCAAGACTTCTGTTAAATGCCAGGGGTGAATGCCAAAACCTGCCGCTCTTG 2610
|||||
Db 30062 GGTCCCAAGACTTCTGTTAAATGCCAGGGGTGAATGCCAAAACCTGCCGCTCTTG 30121
2611 ATGCACACAGTTTAAAGACATCGCAGAAATAGCAGCCCTGTCAAGACAGAGCTCAGAGT 2670
|||||
Db 30122 ATGCACACAGTTTAAAGACATCGCAGAAATAGCAGCCCTGTCAAGACAGAGCTCAGAGT 30181
2671 ATTTCGGGAAATGCTGCAGAAATGCCAAACAGCTTTATGATTTTCACACCTCTTTTGA 2730
|||||
Db 30182 ATTTCGGGAAATGCTGCAGAAATGCCAAACAGCTTTATGATTTTCACACCTCTTTTGA 30241
2731 GAAGTCGTATCAAAAGGAAAGGAAAGGAAAGTGAACAGATGCTGTTTCTTATCCAT 2790
|||||
Db 30242 GAAGTCGTATCAAAAGGAAAGGAAAGGAAAGTGAACAGATGCTGTTTCTTATCCAT 30301
2791 GCCTGTACTTTTCAGCGGCTCCTTGCCAGACATCATAGTCAATTAATTTATGTTG 2850
|||||
Db 30302 GCCTGTACTTTTCAGCGGCTCCTTGCCAGACATCATAGTCAATTAATTTATGTTG 30361
2851 CATTTTCATCTTCTTCCCAATGCTCTTAATGATTTGACGTGACCAAGAACAGATTC 2910
|||||
Db 30362 CATTTTCATCTTCTTCCCAATGCTCTTAATGATTTGACGTGACCAAGAACAGATTC 30421
2911 TCTGTGAAGTGCAGATTAAGCATGATGATGATGATGATGATGATGATGATGATGATG 2970
|||||
Db 30422 TCTGTGAAGTGCAGATTAAGCATGATGATGATGATGATGATGATGATGATGATG 30481
2971 TGACCCGTCTATGCCGGCTTAGCATGTTCTTTTAAATGAGTTTTCAGATCAGGT 3030
|||||
Db 30482 TGACCCGTCTATGCCGGCTTAGCATGTTCTTTTAAATGAGTTTTCAGATCAGGT 30541
3031 AAAGTTCTACAGAGATTTACAGAGGTTAGAACTTTACTATCTTCTTACATCTCATTT 3090
|||||
Db 30542 AAAGTTCTACAGAGATTTACAGAGGTTAGAACTTTACTATCTTCTTACATCTCATTT 30601
3091 TAGAAGGAATATGCTTAAGCCCTGCGATGACGCTCAGGAGGAGGAAAGAGCAGCACAA 3150
|||||
Db 30602 TAGAAGGAATATGCTTAAGCCCTGCGATGACGCTCAGGAGGAGGAAAGAGCAGCACAA 30661
3151 GAAAGCTACATTTTAAACAGTCTTGTATCTAGTGCACATTAATAACAGTCTTAATT 3210
|||||
Db 30662 GAAAGCTACATTTTAAACAGTCTTGTATCTAGTGCACATTAATAACAGTCTTAATT 30721
3211 GCACCTTATACCATGCTGCTGCTGCTCCAAATCTGGCTTGGCTGTGCTGCTGCA 3270
|||||
Db 30722 GCACCTTATACCATGCTGCTGCTGCTCCAAATCTGGCTTGGCTGTGCTGCTGCA 30781
3271 CGCTTGAAGTGAATTTTGTGAGAAATCATTTTGAACCTTTCTTACAAAGAGGCT 3330
|||||
Db 30782 CGCTTGAAGTGAATTTTGTGAGAAATCATTTTGAACCTTTCTTACAAAGAGGCT 30841
3331 TCTGGAACACTGAGAGAAACATCTCTTGGCCATTCCTGACCAAGTTCTCTTACCAATT 3390
|||||

Db 30842 TCTGGACACTGAGAGAAACATCTTTGCCATTCTGCACGAGTTCCTCTACACATTT 30901
Qy 3391 TTCTCAGCTGCATCTTCTGCTGCTGCTCTAGAGAAATTTCAATGAGCCTTCTACT 3450
Db 30902 TTCTCAGCTGCATCTTCTGCTGCTGCTCTAGAGAAATTTCAATGAGCCTTCTACT 30961
Qy 3451 ACTAATTCAGACAGTCTTCCCTGAAAAAAGTGTGAGTCTTCTTAATGACCTTAACATA 3510
Db 30962 ACTAATTCAGACAGTCTTCCCTGAAAAAAGTGTGAGTCTTCTTAATGACCTTAACATA 31021
Qy 3511 TGTAGCATATCTATTAATTTCTGTTGCTCAATTAATGATTTTAAAGCAATGATTTG 3570
Db 31022 TGTAGCATATCTATTAATTTCTGTTGCTCAATTAATGATTTTAAAGCAATGATTTG 31081
Qy 3571 CGTTTGGCAAAAGTTAATGATGAGAGCTTTAGAAATCTCAATTTTGGACATATTA 3630
Db 31082 CGTTTGGCAAAAGTTAATGATGAGAGCTTTAGAAATCTCAATTTTGGACATATTA 31141
Qy 3631 GTCTCCTAATATCAGAGATCCCTAAGTCCAGCTGAGTCTTACAGAGTTTTCAGACTT 3690
Db 31142 GTCTCCTAATATCAGAGATCCCTAAGTCCAGCTGAGTCTTACAGAGTTTTCAGACTT 31201
Qy 3691 CCTGCTTCTCAGCTCTTATATCTTAATGACACACAGCATATCTCTTGAATATACAC 3750
Db 31202 CCTGCTTCTCAGCTCTTATATCTTAATGACACACAGCATATCTCTTGAATATACAC 31261
Qy 3751 TAATGGCAGTGAAGCGAGATGACACACCTGAGCCCTGAGGAGACAGAGTGAAGT 3810
Db 31262 TAATGGCAGTGAAGCGAGATGACACACCTGAGCCCTGAGGAGACAGAGTGAAGT 31321
Qy 3811 TTGTCTCTATTAACAAAAAGAAAAAGAAATACAACTTAAGCTCA 3857
Db 31322 TTGTCTCTATTAACAAAAAGAAAAAGAAATACAACTTAAGCTCA 31368

RESULT 4
US-09-880-107-2225
Sequence 2225, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2225
LENGTH: 35641
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L76568
US-09-880-107-2225

Query Match 47.4%; Score 1829.4; DB 10; Length 35641;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2011 ACTGACACTGGAAGCCGTGGCCAGAGAACAGATGTACACAGCAAGCATAGTTGTG 2070
Db 29522 ACTTACTTTTCTCTGAGTGGCCAGAGAACAGATGTACACAGCAAGCATAGTTGTG 29581
Qy 2071 GATATGCGTGAATTTGCAAGTGAAGCTTCAATCTCTGATCATGTGGGGCATTTGACAT 2130
Db 29582 GATATGCGTGAATTTGCAAGTGAAGCTTCAATCTCTGATCATGTGGGGCATTTGACAT 29641

Qy 2131 GAACCCGTGACCTTTAGAGGTTGAGATTACATCTCACCAGAAATGTCGTTGAGAGCC 2190
Db 29642 GAACCCGTGACCTTTAGAGGTTGAGATTACATCTCACCAGAAATGTCGTTGAGAGCC 29701
Qy 2191 AAGAGTATAGTATTTAATCGGCTCTTAAATACGGCGGCTCTACAGCCAGTGCATC 2250
Db 29702 AAGAGTATAGTATTTAATCGGCTCTTAAATACGGCGGCTCTACAGCCAGTGCATC 29761
Qy 2251 TCCATGTCCGCTACTACAAAGGTCGCCGTCTTGATTTAGTTTGAACCTTAAGAGCT 2310
Db 29762 TCCATGTCCGCTACTACAAAGGTCGCCGTCTTGATTTAGTTTGAACCTTAAGAGCT 29821
Qy 2311 TTCTCTCAGCTTCCGAGGTGCTTTTCAAGAGATCTCCAGCAATGACATTAAGTTC 2370
Db 29822 TTCTCTCAGCTTCCGAGGTGCTTTTCAAGAGATCTCCAGCAATGACATTAAGTTC 29881
Qy 2371 AAACCTACTCTTCTTACACTTCACTTCCCAAGCATCGGAATCTCTGAGCCCTCTCC 2430
Db 29882 AAACCTACTCTTCTTACACTTCACTTCCCAAGCATCGGAATCTCTGAGCCCTCTCC 29941
Qy 2431 CATGCAAGCGGAGTTGTTTGAAGAGCTGAACAAGCAAGCAGCCTGATGCGGCG 2490
Db 29942 CATGCAAGCGGAGTTGTTTGAAGAGCTGAACAAGCAAGCAGCCTGATGCGGCG 30001
Qy 2491 ACAGCACTGGCCATTTACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCT 2550
Db 30002 ACAGCACTGGCCATTTACAGCAGATTTCTGAACCCCTTCCGAGTCAAGAGATTAATCT 30061
Qy 2551 GGTCCCAAGACTCTTCTTTAAAAATGCCAGGGGTGAATGCCAAAAATGCCCTCTTG 2610
Db 30062 GGTCCCAAGACTCTTCTTTAAAAATGCCAGGGGTGAATGCCAAAAATGCCCTCTTG 30121
Qy 2611 ATGCACACAGTTTAAGAACATCGCAGAAATTAGACGCCCTGTCAACAGAGCTACAGAGT 2670
Db 30122 ATGCACACAGTTTAAGAACATCGCAGAAATTAGACGCCCTGTCAACAGAGCTACAGAGT 30181
Qy 2671 ATTCCTGGGGAATCTGCAAAATGCCAAACAGCTTTATGATTTCAATTCACACCTTTTGA 2730
Db 30182 ATTCCTGGGGAATCTGCAAAATGCCAAACAGCTTTATGATTTCAATTCACACCTTTTGA 30241
Qy 2731 GAAGTGTATCAAAAGGAAAAAGGAAAAAGTGAACGTGATGCTGCTTTCTTATCCAT 2790
Db 30242 GAAGTGTATCAAAAGGAAAAAGGAAAAAGTGAACGTGATGCTGCTTTCTTATCCAT 30301
Qy 2791 GCCTGTACTTTGAGCGGCTCTTGGCAGACATATAGTCATTAATTAATTTGTTG 2850
Db 30302 GCCTGTACTTTGAGCGGCTCTTGGCAGACATATAGTCATTAATTAATTTGTTG 30361
Qy 2851 CTATTTCAATCTTTTCCAAATGCTTTAATGATTTGACGGTGAACCCAGGATTC 2910
Db 30362 CTATTTCAATCTTTTCCAAATGCTTTAATGATTTGACGGTGAACCCAGGATTC 30421
Qy 2911 TCTCTGACCTGACAGTATGAGCATCTTGAACCTGCTGCTCTCTTTTCCCTCC 2970
Db 30422 TCTCTGACCTGACAGTATGAGCATCTTGAACCTGCTGCTCTCTTTTCCCTCC 30481
Qy 2971 TGCACGCTATGCGGGCTTACAGATGTTCTTTTAAATGAGGTGTGTCAGATCAGT 3030
Db 30482 TGCACGCTATGCGGGCTTACAGATGTTCTTTTAAATGAGGTGTGTCAGATCAGT 30541
Qy 3031 AAAGTTCTACAGTATTAACAGAGGTACAACTTACCTGATCTTAACAGATCTCAT 3090
Db 30542 AAAGTTCTACAGTATTAACAGAGGTACAACTTACCTGATCTTAACAGATCTCAT 30601
Qy 3091 TAGAAAGGAATATGCTTAAGCCTGGCATGACGGGTGAGGAGGAAAAAGAGCAGCA 3150
Db 30602 TAGAAAGGAATATGCTTAAGCCTGGCATGACGGGTGAGGAGGAAAAAGAGCAGCA 30661
Qy 3151 GAAAGCTACATTTTAAAGCTCTTGTATCTATGTCACAAATTAATTAAGAGTCTTAAT 3210
Db 30662 GAAAGCTACATTTTAAAGCTCTTGTATCTATGTCACAAATTAATTAAGAGTCTTAAT 30721

OY	3211	GCACCTATACCACTGCTCGTGGGCTCCCAATCGTGGCTTGGCTGGCTGGCTGGA	3270
Db	30722	GCACCTATACCACTGCTCGTGGGCTCCCAATCGTGGCTTGGCTGGCTGGCTGGA	30781
OY	3271	CGCTTGAACGTATGTTTGCTAGGAAATCATGTTCTGACCCCTTGTCTACAAAGAGCCT	3330
Db	30782	CGCTTGAACGTATGTTTGCTAGGAAATCATGTTCTGACCCCTTGTCTACAAAGAGCCT	30841
OY	3331	TCCTGGACACCTGAGAGAAACATCTCTTGGCATTTCTGACACAGTTTCTCTACACATTT	3390
Db	30842	TCCTGGACACCTGAGAGAAACATCTCTTGGCATTTCTGACACAGTTTCTCTACACATTT	30901
OY	3391	TTCTTCACCTCCATCTCTCTGCTGTGCTCTAGAGAAATTTCAATGAGAGCCTTCTCTACT	3450
Db	30902	TTCTTCACCTCCATCTCTCTGCTGTGCTCTAGAGAAATTTCAATGAGAGCCTTCTCTACT	30961
OY	3451	ACTAATTTAAGACAGCTCTCCCTCAAAAACCTGGTTGACTAGTCTCTTAATGACCTTAACATA	3510
Db	30962	ACTAATTTAAGACAGCTCTCCCTCAAAAACCTGGTTGACTAGTCTCTTAATGACCTTAACATA	31021
OY	3511	TGTACACATATCTATTAATTTCAATTTGTCCTCAATTTGTAATTTTAAGCAAAATGAAATTAAC	3570
Db	31022	TGTACACATATCTATTAATTTCAATTTGTCCTCAATTTGTAATTTTAAGCAAAATGAAATTAAC	31081
OY	3571	CTGTTTGGCAAAAGTTAATGATGAGAGAGCTCTTAGAATTTCTCAATTTTTGGACATATTTCA	3630
Db	31082	CTGTTTGGCAAAAGTTAATGATGAGAGAGCTCTTAGAATTTCTCAATTTTTGGACATATTTCA	31141
OY	3631	GTCCTCTAATATCAAGAGATCCCTAGTCCAGCTGGCTGTGTTACAGATTTTTTTCAGACTT	3690
Db	31142	GTCCTCTAATATCAAGAGATCCCTAGTCCAGCTGGCTGTGTTACAGATTTTTTTCAGACTT	31201
OY	3691	CCTGCTTTCAGAGCTCTTATATCTTAATCTTAAGACACACACATCATATTCCTCTTGAATATCAACC	3750
Db	31202	CCTGCTTTCAGAGCTCTTATATCTTAATCTTAAGACACACACATCATATTCCTCTTGAATATCAACC	31261
OY	3751	TAAATGGCAGTGAGCCGAGATCGCACACACCTGACCCCTGCTGGCGACAGAGTAGACT	3810
Db	31262	TAAATGGCAGTGAGCCGAGATCGCACACACCTGACCCCTGCTGGCGACAGAGTAGACT	31321
OY	3811	TTTGCTCATTACAAAAGAAAAGAAAAGAAATACACCTAGACATCA 3857	
Db	31322	TTTGCTCATTACAAAAGAAAAGAAAAGAAATACACCTAGACATCA 31368	
RESULT 5			
US-09-867-701-10472/c			
Sequence 10472, Application US/09867701			
Patent No. US20020132237A1			
GENERAL INFORMATION:			
APPLICANT: Agiate, Paul A.			
APPLICANT: Jones, Robert			
APPLICANT: Harlocker, Susan L.			
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER			
FILE REFERENCE: 210121.497			
CURRENT APPLICATION NUMBER: US/09/867,701			
CURRENT FILING DATE: 2001-05-29			
NUMBER OF SEQ ID NOS: 10912			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 10472			
LENGTH: 601			
TYPE: DNA			
ORGANISM: Homo sapien			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (1)...(601)			
OTHER INFORMATION: n - A..T..C or G			
US-09-867-701-10472			

Query Match	6.6%;	Score 256.2;	DB 10;	Length 601;
Best Local Similarity	97.7%;	Pred. No. 1.2e-53;		
Matches 258;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;

Qy	1561	GATTATAGCAGTATAGCCGAAACCTGCCCCGAGAAATTAACCATCAAGAAATTTGATGTA	1620
Db	601	GAAATACGACGTATAGCCCGAGNACCTGCCCCGAAAGAAATTAACCATGAAATATGATGTA	542
Qy	1621	AATTTGTCATCGATGTGCTGTTTCGGGAAATCTGTAAAGAACCCCTCACTTCATTCATCCG	1680
Db	541	AATTTGTCATCGATGTGCTGTTTCGGGAAATCTGTAAAGAACCCCTCACTTCATTCATCCG	482
Qy	1681	CTTCTGGTGTGCACGACCCCTATGCTGTGTGACAAAGGTACTCAATCAAGTGGAGCCAGA	1740
Db	481	CTTCTGGTGTGCACGACCCCTATGCTGTGTGACAAAGGTACTCAATCAAGTGGAGCCAGA	422
Qy	1741	TACGTGGTGTCTTTATGACGACGAGACTAACTTTGTTGGCAGACTTGAAAATTCAGGGCG	1800
Db	421	TACGTGGTGTCTTTATGACGACGAGACTAACTTTGTTGGCAGACTTGAAAATTCAGGGCG	362
Qy	1801	AGTAGGCTGGGAAACCTCTGAGG	1824
Db	361	AGTAGGCTGGGAAACCTCTGAGG	338

```

: RESULT 6
: US-09-919-497-14/C
: Sequence 14, Application US/09919497
: Patent No. US2002010662A1
: GENERAL INFORMATION:
: APPLICANT: Mutter, George L.
: TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
: FILE REFERENCE: B0801/7225
: CURRENT APPLICATION NUMBER: US/09/919,497
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/221,735
: PRIOR FILING DATE: 2000-07-31
: NUMBER OF SEQ ID NOS: 100
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 14
: LENGTH: 3857
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-919-497-14

```

Query Match	4.08;	Score 154.6;	DB 10;	Length 3857;
Best Local Similarity	82.08;	Pred. No. 3.9e-28;		
Matches 178; Conservative	0;	Mismatches 39;	Indels 0;	Gaps 0;

QY	2941	AAATTCGCTGCGCCCTGCCTCTTTTTCGTCCTCCGACGCTGTATGCGGGCTTAGCAATGTT	3000
Db	3157	AGCTTCTCTTGCCCTGCTCTTTTCCCTCCCTGCACCGCTCAGAGGCTTAGCAATTC	3098
QY	3001	CTTTTAAATAGGTTTGTCAGCATCAGTAAAGTTCTCTACAAGTATATACGAAGTAG	3060
Db	3097	CTTCTCAATATAGATCTGTTAGATCAAGTAAAGTTCTTACCTCTGTATCACTGTAG	3038
QY	3061	AAACTTACCTGATCCTPACAGATCTTCATTAGAAAGAAATATGCTAACCTGGCATGGA	3120
Db	3037	GAACCTTACCTGATCTCGACAAACCTCATTTAAAAAGAAACATGCTTAAACCCGCATAGA	2978
QY	3121	CGGTGACAGGAGGAAAAAGACGACGCAAGAAGCT	3157
Db	2977	CGGTGACAGGAGGAAAAAGACGACGCAAGAAGTT	2941

```

? RESULT 7
? US-09-962-436-306/C
? Sequence 306, Application US/09962436
? Patent No. US20020081301A1
? GENERAL INFORMATION:
? APPLICANT: Soppet, Daniel
? TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening using Sigm
? TITLE OF INVENTION: Sets
? FILE REFERENCE: 689290-75
? CURRENT APPLICATION NUMBER: US/09/962,436

```



```

RESULT 11
US-09-764-877-2510
; Sequence 2510, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2510
; LENGTH: 3705
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2510

Query Match          1.8%; Score 69.2; DB 10; Length 3705;
Best Local Similarity 85.6%; Pred. No. 4.5e-07;
Matches 77; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY   3755 TGGCAGTAGGCCGAGATCGACCACTGCACCCTCCGTGGGGCGACAGATGACTTGT 3814
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    1900 TTGCAGTAGGCCAGATCGCACCTGCACCTCCAGCTCGGCGAACAGTAGACTCTGT 1959

OY   3815 CTCTATTACAAAAGAAAGAAAGAATA 3844
      ||| || | | | | | | | | | | | | | |
Db    1960 CTCAAAAAAGAAAGAAAGAAAGATGTA 1989


RESULT 12
US-09-969-708-79
; Sequence 79, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 79
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-79

Query Match          1.8%; Score 69.2; DB 10; Length 145831;
Best Local Similarity 81.6%; Pred. No. 3.4e-06;
Matches 80; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY   3755 TGGCAGTAGGCCGAGATCGACCACTGCACCCTCCGTGGGGCGACAGATGACTTGT 3814
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    38704 TTGCAGTAGGCAGAGATCGCACCACTGCACCTCCAGCTCGGCTGGAGACAGAGACTCTGT 38763

OY   3815 CTCTATTACAAAAGAAAGAAAGAATAATACACTTAA 3852
      ||| || | | | | | | | | | | | | | | ||| ||
Db    38764 CTCAAAAAAGAAAGAAAGAAAGAAAGAAACTGAA 38801


RESULT 13
US-09-954-456-2116
; Sequence 2116, Application US/09954456

```

```

Patent No. 020020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2116
LENGTH: 145831
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-2116

Query Match      1.8%; Score 69.2; DB 10; Length 145831;
Best Local Similarity 81.6%; Pred. No. 3.4e-06;
Matches 80; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy   3755 TGGCGATGAGCCGAGATGCACACCCTGCCCTGGGCGACAGTGAGACTTTGT 3814
      | ||||| ||| |||||||||||||||| ||| ||||| ||||| ||||| |||
Db   38704 TTGCGATTAGCAGAGATGCACCTGCATCTCCAGCTGGGTGACACGACGACTCTGT 38763
Oy   3815 CTCTATTACAAGAAAGAAAAGAATAATACACCTAA 3852
      ||| | ||| ||| ||||| ||||| ||| |||
Db   38764 CTCAAAAAAGAAAAAGAAAAGAAAAGAAACTGAA 38801

RESULT 14
US-09-860-670-263
Sequence 263, Application us/09860670
Patent No. US20020165137A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA127P1
CURRENT APPLICATION NUMBER: US/09/860,670
CURRENT FILING DATE: 2001-05-21
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 289
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 263
LENGTH: 24218
TYPE: DNA
ORGANISM: Homo sapiens
US-09-860-670-263

Query Match      1.8%; Score 69; DB 9; Length 24218;
Best Local Similarity 83.9%; Pred. No. 1.4e-06;
Matches 78; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy   3755 TGGCAGTGAAGCCGAGATGCACACCCTGCCCTGGGCGACAGTAGAGACTTTGT 3814

```

Db 6215 TTGCAGTGAGCCGAGATCGCACCCTGCGACTCTCGGCGACAGATGAGACTCCGT 6274

Qy 3815 CTCTATTACAAAAGAAAAGAAAAGAAATACA 3847

Db 6275 CTCAAAACAAAAAAGAAAAAAGAAAAA 6307

RESULT 15

US-09-764-410

; Sequence 410, Application US/09764878

; Patent No. US2002090615A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA121

; CURRENT APPLICATION NUMBER: US/09/764,878

; PRIOR APPLICATION DATE: 2001-01-17

; PRIOR APPLICATION data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 428

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 410

; LENGTH: 323

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-878-410

Query Match

Best Local Similarity 1.8%; Score 68; DB 10; Length 323;

Matches 77; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 3755 TGGCAGTGAGCCGAGATCGCACCCTGCGCTGGGCGACAGAGTGAAGTTGT 3814

Db 219 TTGCAGTGAGCCGAGATCGCACCCTGCGACTCGGCGACAGAGTGAAGTTGT 278

Qy 3815 CTCTATTACAAAAGAAAAGAAAAGAAATACA 3846

Db 279 CTCAAAACAAAAAAGAAAAAAGAGA 310

Search completed: November 26, 2002, 23:47:25

Job time : 1153.71 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

GenEmb1: *

- 1: gb_ba: *
- 2: gb_bg: *
- 3: gb_in: *
- 4: gb_om: *
- 5: gb_ov: *
- 6: gb_pat: *
- 7: gb_ph: *
- 8: gb_pl: *
- 9: gb_pr: *
- 10: gb_ro: *
- 11: gb_ses: *
- 12: gb_sy: *
- 13: gb_un: *
- 14: gb_vl: *
- 15: em_ba: *
- 16: em_bu: *
- 17: em_fun: *
- 18: em_in: *
- 19: em_mu: *
- 20: em_om: *
- 21: em_or: *
- 22: em_ov: *
- 23: em_pat: *
- 24: em_ph: *
- 25: em_pl: *
- 26: em_ro: *
- 27: em_ses: *
- 28: em_un: *
- 29: em_vl: *
- 30: em_bg_bum: *
- 31: em_bg_iny: *
- 32: em_bg_other: *
- 33: em_bg_mus: *
- 34: em_bg_pin: *
- 35: em_bg_rod: *
- 36: em_bg_tam: *
- 37: em_bg_vrt: *
- 38: em_sy: *
- 39: em_hgo_bum: *
- 40: em_hgo_mus: *
- 41: em_hgo_other: *

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	672	100.0	672	9	HMGGA7A08	M93036 Human (clon
2	672	100.0	171887	9	AC079775	AC079775 Homo sapi
3	461.6	68.7	621	6	AX197746	AX197746 Sequence
4	461.6	68.7	621	6	AX208351	AX208351 Sequence
5	451.4	67.2	1545	6	AX014895	AX014895 Sequence
6	449.8	66.9	1565	9	AX026585	AX026585 Homo sapl
7	449.4	66.9	1504	9	HKSKA	X14758 Human mRNA
8	449.4	66.9	1504	9	HMGKSA	M3235 Human adeno
9	447.8	66.6	1457	9	HMGKSA	M3231 Human (clon
10	442.4	66.0	637	6	AX395799	AX395799 Sequence
11	442.4	65.8	481	6	AX397344	AX397344 Sequence
12	442.2	65.8	545	6	AX198687	AX198687 Sequence
13	439.4	65.4	1511	9	HUMMAK514	M26481 Homo sapien
14	439	65.3	439	6	AX209414	AX209414 Sequence
15	438.4	65.2	1372	9	HDMSCAE	M32306 Human epith
16	415.4	61.8	425	6	AX193024	AX193024 Sequence
17	362.4	53.9	1458	9	BC014785	BC014785 Homo sapl
18	342.6	51.0	173814	9	AC093840	AC093840 Homo sapl
19	335.8	50.0	507	6	AX209701	AX209701 Sequence
20	333.8	49.7	555	6	AX395792	AX395792 Sequence
21	326	48.5	326	6	AX966230	AX966230 Sequence
22	269	40.0	269	6	AX192570	AX192570 Sequence
23	234	34.8	235	11	G43776	G43776 WIAF -3376-S
24	222.8	33.2	272	6	AX182073	AX182073 Sequence
25	222	33.0	353	6	AX351467	AX351467 Sequence
26	221.2	32.9	223	6	AX340829	AX340829 Sequence
27	221.2	32.9	223	6	AX341705	AX341705 Sequence
28	220.2	32.8	223	6	AX340713	AX340713 Sequence
29	214	31.8	214	6	106778	106778 Sequence 4
30	199.2	29.6	201	6	AX341941	AX341941 Sequence
31	185.4	27.6	691	6	AX351469	AX351469 Sequence
32	183.6	27.3	238	6	AX340174	AX340174 Sequence
33	183.4	27.3	554	6	AX192846	AX192846 Sequence
34	182.4	27.1	205	6	AX341377	AX341377 Sequence
35	178	26.5	178	6	AX392711	AX392711 Sequence
36	171.6	25.5	462	6	AX339773	AX339773 Sequence
37	145	21.6	160	6	AX197888	AX197888 Sequence
38	145	21.6	160	6	AX208485	AX208485 Sequence
39	112	16.7	191914	2	AC104671	AC104671 Mus muscu
40	107	15.9	174433	2	AC101909	AC101909 Mus muscu
41	96.6	14.4	84604	10	AL656763	AL656763 Mouse DNA
42	93.8	14.0	1492	10	BC005618	BC005618 Mus muscu
43	91.2	13.6	1710	6	AX305882	AX305882 Sequence
44	91.2	13.6	1710	10	M0SECP	M76124 Mus muscunu
45	86.2	12.8	1524	10	RNEG314H	A7001044 Rattus no

ALIGNMENTS

	RESULT	1
HUMGA7A08	Locus	HUMGA7A08
DEFINITION		HUMGA7A08 672 bp DNA]linear PRI_23-MAY-1996
ACCESSION		Human (clone Z1726) carcinoma-associated antigen GA733-2 (GA733-2)
VERSION		mRNA, exon 9 and complete cds.
KEYWORDS		M93036 M93036.1 GI:182904 carcinoma-associated antigen GA733-2. 8 Of 8
SOURCE		Homo sapiens (tissue library: UNB7) lymphoma DNA.
ORGANISM		Homo sapiens Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE		1 (styes) Szala,S., Froehlich,M., Scollon,M., Kasai,Y., Stepiewski,Z.,
AUTHORS		

Koprowski, H. and Linnenbach, A. J.
Molecular cloning of cDNA for the carcinoma-associated antigen
CA733-2
Proc. Natl. Acad. Sci. U.S.A. 8, 3542-3546 (1990)
2 (bases 1 to 672)
Linnenbach, A. J., Seng, B. A., Wu, S., Robbins, S., Scollon, M.,
Pyrce, J. J., Druck, T. and Huebner, K.
Retroposition in a family of carcinoma-associated antigen genes
Mol. Cell. Biol. 13 (3), 1507-1515 (1993)
JOURNAL
MEDLINE
93180797
PUBMED
8382772

FEATURES
source
1. 672
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="iv"
/cell_type="B-cell"
/tissue_type="lymphoma"
/feature_id="LN87"
join(M93029,1:35, .573,M93030,1:1, .759,M93031,1:1, .355,
M93032,1:1, .367,M93033,1:1, .596,M93034,1:1, .568,
M93035,1:1, .336,1, .606)
/gene="CA733-2"
join(M93029,1:269, .522,M93030,1:83, .190,M93030,1:428, .668,
M93031,1:30, .95,M93032,1:163, .226,M93033,1:304, .405,
M93034,1:265, .465,M93035,1:95, .139,160, .606)
/gene="GA733-2"
join(M93029,1:447, .532,M93030,1:83, .190,M93030,1:428, .668,
M93031,1:30, .95,M93032,1:163, .226,M93033,1:304, .405,
M93034,1:265, .465,M93035,1:95, .139,160, .201)
/gene="GA733-2"
/codon_start=1
/product="carcinoma-associated antigen GA733-2"
/protein_id="AAB00775.1"
/db_xref="GI:182906"
/translation="MAPPOVLAFLGLLAATRTFAAODEVCENYKKAIVNCFVNNR
OCQCTSVGAQNTVICSILAAKCLVMAKENGSKLRARPEGALQNNGLDPPDCDES
GLFAKOCNGSTCVNTAGVRTDKETISERVRYWIIIEIKHRAKREPYDESK
SLRTALOKETITRYOLOPKETISILYENNVITIDLVONSSCKTQNDVDAVAYFEK
DYKGSLEFHSKKMLVNGEOLDLDPGOTLIYVDEKAEFEFMQGLKAGVIAVIVVV
IAVAGIVAVLVSRKKRAKYEKAEIKEMGHRELNA"
order(M93035,1:140, .336,1, .159)
/gene="GA733-2"
/number=8
160, .606
/gene="GA733-2"
/number=9
500, .505
/gene="GA733-2"
518
/gene="GA733-2"
588, .593
/gene="GA733-2"
606
polyA_site
/gene="GA733-2"
BASE COUNT 222 a 85 c 100 g 265 t
ORIGIN
Query Match 100.0%; Score 672; DB 9; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.7e-101;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CATAGGAACTCATGTCATATATATATTTGAAGATTATAGAA :AAGGAANTAGCAA 240
|||||
DB 181 CATAGGAACTCATGTCATATATATATTTGAAGATTATAGAA :AAGGAANTAGCAA 240
QY 241 ATGACACAAATATACAAATGTGTGTCGTCGAGCAAGACATCTTGAAGTCATAGATT 300
|||||
DB 241 ATGACACAAATATACAAATGTGTGTCGTCGAGCAAGACATCTTGAAGTCATAGATT 300
QY 301 TGTAGTTTACATCATATATTTTGTAAATAGTGAACCTGTACTCAAAATATAGACAGTT 360
|||||
DB 301 TGTAGTTTACATCATATATTTTGTAAATAGTGAACCTGTACTCAAAATATAGACAGTT 360
QY 361 GAATGCGCTTATCCATGTCATGTAATTTGACCCAGAGTCTTATATATGATGATATAT 420
|||||
DB 361 GAATGCGCTTATCCATGTCATGTAATTTGACCCAGAGTCTTATATATGATGATATAT 420
QY 421 GTAAATCCAGAACTTGACCTCCATCGTTAAATTTATATGTATGTAACATTCATATGTGT 480
|||||
DB 421 GTAAATCCAGAACTTGACCTCCATCGTTAAATTTATATGTATGTAACATTCATATGTGT 480
QY 481 GCATTAAATATGCTTCCACAGTAAATTCGAAAACTGATTTGTGATTGAAAGCTGCTT 540
|||||
DB 481 GCATTAAATATGCTTCCACAGTAAATTCGAAAACTGATTTGTGATTGAAAGCTGCTT 540
QY 541 TCTATTACTTGAGCTTGTACATACATCTTTTATAGCATATGAAATTAACAATTTT 600
|||||
DB 541 TCTATTACTTGAGCTTGTACATACATCTTTTATAGCATATGAAATTAACAATTTT 600
QY 601 AAACGTAATTTCTTAACCTTGACATTTCAAAATTCCTCTCTCTCTCTCTCTCTCTCT 660
|||||
DB 601 AAACGTAATTTCTTAACCTTGACATTTCAAAATTCCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 TTTTCTTTGAGA 672
|||||
DB 661 TTTTCTTTGAGA 672
|||||

RESULT 2
AC079775 171987 bp DNA linear PRI 29-MAY-2002
LOCUS
DEFINITION
AC079775 Homo sapiens BAC clone RP11-295P2 from 2, complete sequence.
AC079775
AC079775.6 GI:19848453
VERSION
KEYWORDS
SOURCE
Homo sapiens.
Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Sulston, J. E. and Waterston, R.
TITLE
JOURNAL
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE
99063792
PUBMED
9847074
REFERENCE
2 (bases 1 to 171987)
AUTHORS
Belter, E., Haekenson, W., Doeber, A. and Elliott, G.
TITLE
JOURNAL
The sequence of Homo sapiens BAC clone RP11-295P2
Unpublished (2001)
REFERENCE
3 (bases 1 to 171987)
AUTHORS
Waterston, R. H.
TITLE
JOURNAL
Direct Submision
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 171987)
AUTHORS
Waterston, R. H.
TITLE
JOURNAL
Direct Submision
Submitted (30-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 171987)
AUTHORS
Waterston, R.

TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 30, 2002 this sequence version replaced g1.16924154.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0295P02

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McHersom, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPI1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RPI1-261E8; the clone sequenced to the right is RPI1-436K12. Actual start of this clone is at base position 1 of RPI1-295P2; actual end is at base position 171987 of RPI1-295P2.

FEATURES

source

1. 171987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RPI1-295P2"
/clone_id="RPI1-11"
496. 539
/rpt_family="MERL_type"
540. 840
/rpt_family="Alu"
841. 1002
/rpt_family="MERL_type"
1003. 1306
/rpt_family="Alu"
1307. 1369
/rpt_family="MERL_type"
1703. 1951
/note="match to EST BE504918 (NID:g9707326) hz32e08.x1"
1770. 2076
/rpt_family="Alu"
2455. 2832
/note="similar to Homo sapiens EST BF819288 (NID:g12157087)"
2612. 2687
repeat_region

/rpt_family="MIR"
2688. 2989
/rpt_family="Alu"
2990. 3052
repeat_region
/rpt_family="MIR"
3268. 3287
repeat_region
/rpt_family="(A)n"
3305. 3412
/rpt_family="L1"
3577. 3886
/rpt_family="Alu"
4004. 4321
repeat_region
/rpt_family="Alu"
4322. 4626
repeat_region
/rpt_family="ERV1"
4629. 4703
/rpt_family="Alu"
4669. 4695
misc_feature
/note="match to EST BG982215 (NID:g14384950)"
5895. 5994
/rpt_family="MIR"
6612. 6975
repeat_region
/rpt_family="MALR"
7608. 8109
/note="match to EST BF105129 (NID:g10887655)"
7625. 7727
/rpt_family="MIR"
7735. 7876
/rpt_family="MERL_type"
8217. 8547
repeat_region
/rpt_family="L2"
8984. 9245
repeat_region
/rpt_family="Alu"
9251. 9279
/rpt_family="(CAA)n"
9283. 9334
repeat_region
/rpt_family="Alu"
9335. 9640
repeat_region
/rpt_family="Alu"
9641. 9800
repeat_region
/rpt_family="Alu"
9802. 10082
repeat_region
/rpt_family="Alu"
10088. 10118
repeat_region
/rpt_family="(TAA)n"
10385. 10465
repeat_region
/rpt_family="MIR"
10667. 10961
repeat_region
/rpt_family="Alu"
11074. 11530
repeat_region
/rpt_family="MALR"
11730. 12034
repeat_region
/rpt_family="Alu"
12039. 12322
repeat_region
/rpt_family="Alu"
12372. 12408
repeat_region
/rpt_family="(TCCC)n"
12461. 12651
repeat_region
/rpt_family="MIR"
12747. 13036
repeat_region
/rpt_family="Alu"
13097. 13413
repeat_region
/rpt_family="Alu"
13445. 13730
repeat_region
/rpt_family="Alu"
13841. 13968
repeat_region
/rpt_family="MIR"
13969. 14260
repeat_region
/rpt_family="Alu"
14261. 14368
repeat_region
/rpt_family="MIR"
14412. 14720
repeat_region
/rpt_family="Alu"

RESULT 3					
AXI97746/c	AXI97746	621 bp	DNA	linear	PAT 29-AUG-2001
LOCUS	Sequence 201 from Patent WO0151513.				
DEFINITION	AXI97746				
ACCESSION	AXI97746				
VERSION	AXI97746.1	GI:15388061			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	I (bases 1 to 621)				
AUTHORS	Algate,P.A.				
TITLE	Ovarian tumor-associated sequences				
JOURNAL	Patent: WO 0151513-A 201 19-JUL-2001;				
	CORIXA CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1..621				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	207 a	118 c	82 g	195 t	19 others
ORIGIN					

Query Match	Similarity	68.7%	Score	461.6	DB	6	Length	621
Best Local	Similarity	91.3%	Pred.	1.1e-66				
Matches	525	Conservative	0	Mismatches	44	Indels	6	Gaps
								4
QY	26	AAATAGTCTTATAAATGTGGGAAAAAATTAATCTGTGTCTTCTTAAATTCATTTTAT	85					
Db	569	AAATAGTCTTATAAATGTGGGAAAAAATTTCTGTG---TTCNTTAAATTCATTTTAT	513					
QY	86	TAAATCTTTTTCACAAATGCAAAAGATTTGAAAATATATTAGAAATTTTCTGTGCT	145					
Db	512	TAAATCTTTTTCACAAATGCAAAAGATTTGAAAATATATTAGAAATTTTCTGTGCT	455					
QY	146	TTTTCTGTTCAGATTAAGGAGATGGGTGAGATGATCATAGGGAACTCAATGATTAAT	205					
Db	454	TTTTCTGTTCAGATTAAGGAGAGAGGTGAGATGATCATAGGGAACTCAATGATTAAT	396					
QY	206	ATAATTTGAGATTATAGAGAGGGAATATGCAAAATGAGACAAATTAACAATGTGTGT	265					
Db	395	ATACTTTGAGATTATAGAGAGGGAATATGCAAAATGAGACAAATTAACAATGTGTGT	336					
QY	266	GCGTGGGACGAAGACATCTTTGAAGCTGATGAGTTTGTGATTAACTATCATATTTTGT	325					
Db	335	GCGTGGGACGAAGACATCTTTGAAGCTGATGAGTTTGTGATTAACTATCATATTTTGT	276					
QY	326	AATAGTGAAGACCTGACTCAAAATTTAAGACGCTTGAAGACTGGCTTTACCAATCTTGAA	385					
Db	275	AATAGTGAAGACCTGACTCAAAATTTAAGACGCTTGAAGACTGGCTTTACCAATCTTGAA	216					
QY	386	TTTGACCAACAAGTGTCTTATATATGACAGATCTAATGTAAATTCGAACCTTGGACAT	445					
Db	215	TTTGACCAACAAGTGTCTTATATATGACAGATCTAATGTAAATTCGAACCTTGGACAT	156					
QY	446	CGTTAAATTTATTTATGTGTAAATTCGATTAATGTGTGCTTCCACAGTAA	505					
Db	155	CGTTAAATTTATTTATGTGTAAATTCGATTAATGTGTGCTTCCACAGTAA	96					
QY	506	ATCTGAAAAAGTATTTGATTTGAAGACCTGCTTCAATTACTTGATGCTGTGACATA	565					
Db	95	ATCTGAAAAAGTATTTGATTTGAAGACCTGCTTCAATTACTTGATGCTGTGACATA	36					
QY	566	CATACCTTTTATGAGCTATGAATTAACAATTTT	600					
Db	35	CATGCTTTTATGAGCTATGAATTAACAATTTT	1					

RESULT 4

AX208351/c

AX208351/c

621 bp

DNA

linear

PAT 31-AUG-2001

DEFINITION Sequence 191 from Patent WO0157207.
ACCESSION AX208351
VERSION AX208351.1 GI:15422774
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Aligete, P.A. and Mannion, J.
TITLE 1 (bases 1 to 621)
JOURNAL Compositions and methods for the therapy and diagnosis of ovarian
cancer.
PATENT: WO 0157207-A 191 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..621
BASE COUNT /organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
207 a 118 c 82 g 195 t 19 others
Query Match 68.7%; Score 461.6; DB 6; Length 621;
Best Local Similarity 91.3%; Pred. No. 1.1e-66;
Matches 525; Conservative 0; Mismatches 44; Indels 6; Gaps 4;
QY 26 AATAGCTCTAATTAATGGGAAAAAATTAATCTGCTGCTTCAATTAATTTTAT 85
DB 569 AATAGCTCTAATTAATGGGAAAAAATTTCTTG---TTCTTTATTTCTTTTAT 513
QY 86 TAATACATTTTTCAGAAATGAACAAAGATGAAAAATTTAGAAATTTTCTGCT 145
DB 512 TAAGCTATTTTTCANGATG-ACNNAAGATGAAAAATTTATNAG-ATTTTCTGCT 455
QY 146 TTTTCCTTTTCAGATTAAGAGATGGTGATGACATAGGAGACTCAATGATACAT 205
DB 454 TTTTCCTTTTCAGATTAAGAGAGATGGTGATGACATAGGAGACTCAATGATACAT 396
QY 206 AATATTTGAATTTATAGAGAGAGGAAATAGCAATAGACAAATTAACAAATGTGT 265
DB 395 ATACTTTGAATTTATAGAGAGAGGAAATAGCAATAGGAGACTCAATGATACAT 336
QY 266 GCGTGGAGAGAGACATCTTTGAAGTCATAGTTTGTAGTTTAAATCATATATTTGT 325
DB 335 GCGTGGAGAGAGACATCTTTGAAGTCATAGTTTGTAGTTTAAATCATATATTTGT 276
QY 326 AATAGTGAACCTGTACTCAAAATTTAGCAGCTTGAACGCTTTACCAATCTTGAA 385
DB 275 AATAGTGAACCTGTACTCAAAATTTAGCAGCTTGAACGCTTTACCAATCTTGAA 216
QY 386 TTTGACCAAGTGTCTATATATGATGATTAATTAATCCAGACTTGACATCAT 445
DB 215 TTTGACCAAGTGTCTATATATGATGATTAATTAATCCAGACTTGACATCAT 156
QY 446 CGTTAAATTTATTTATGTGTAACATTCAAATGTGCAATTAATTTGCTTCCACAGTAA 505
DB 155 CGTTAAATTTATTTATGTGTAACATTCAAATGTGCAATTAATTTGCTTCCACAGTAA 96
QY 506 ATCTGAATAAAGCTGATTTGATGTAAGAGCTGCTTCAATTAATTAATTTGATGTA 565
DB 95 ATCTGAATAAAGCTGATTTGATGTAAGAGCTGCTTCAATTAATTTGATGTA 36
QY 566 CATACTTTTATGAGCTATGAAATTAATTAATTTT 600
DB 35 CATGCTTTTATGAGCTATGAAATTAATTAATTTT 1
RESULT 5
AX014895 1545 bp DNA Linear PAT 07-SEP-2000
LOCUS AX014895
DEFINITION Sequence 94 from Patent WO9953040.
ACCESSION AX014895
VERSION AX014895.1 GI:10041162
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinemann, B., Rosenthal, A. and
Pilarsky, C.
TITLE Human nucleic acid sequences from ovarian tumor tissue
JOURNAL Patent: WO 9953040-A 94 21-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source location/Qualifiers
1..1545
BASE COUNT /organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
460 a 307 c 369 g 409 t
Query Match 67.2%; Score 451.4; DB 6; Length 1545;
Best Local Similarity 99.8%; Pred. No. 4.1e-65;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 156 TCAGATAAAGGAGATGGGTGAGATGCATAGGAGACTCAATGATTAATTAATTTGAA 215
DB 1077 TGAGATTAAGGAGATGGGTGAGATGCATAGGAGACTCAATGATTAATTAATTTGAA 1136
QY 216 GATTATAGAGAGGAAATAGCAAAATGACACAAATTTCAAAATGTGTGCGTGGACG 275
DB 1137 GATTATAGAGAGGAAATAGCAAAATGACACAAATTTCAAAATGTGTGCGTGGACG 1196
QY 276 AAGACATCTTTGAAGCTGATGATGTTGTAGTTTAACTCAATATTTGTATATAGTAAA 335
DB 1137 AAGACATCTTTGAAGCTGATGATGTTGTAGTTTAACTCAATATTTGTATATAGTAAA 1256
QY 336 CCTGACTCAAAATTAATGAAGAGCTTGAAGAGCTTTCACATCTTGAATTTGACACA 395
DB 1257 CCTGACTCAAAATTAATGAAGAGCTTGAAGAGCTTTCACATCTTGAATTTGACACA 1316
QY 396 AGTGTCTTATATATGACATCTAATATGTAATTCGAACCTTGACATCTGTTAAATTT 455
DB 1317 AGTGTCTTATATATGACATCTAATATGTAATTCGAACCTTGACATCTGTTAAATTT 1376
QY 456 ATTATGTATACATTTCAAAATGTGCAATTAATTTGCTTCCACAGTAAATCTGAAAA 515
DB 1377 ATTATGTATACATTTCAAAATGTGCAATTAATTTGCTTCCACAGTAAATCTGAAAA 1436
QY 516 CTGATTTGTGATTTGAAGCTGCTTCAATTAATTTGATGATCTTGTATACATACATCTTTT 575
DB 1437 CTGATTTGTGATTTGAAGCTGCTTCAATTAATTTGATGATCTTGTATACATACATCTTTT 1496
QY 576 TATGAGCTATGAAATTAATTAATTTAACTGAA 608
DB 1497 TATGAGCTATGAAATTAATTAATTTAACTGAA 1529
RESULT 6
AK026585 1565 bp mRNA Linear PRI 29-SEP-2000
LOCUS AK026585
DEFINITION Homo sapiens cDNA: FLJ22932 fls, clone KAT07515, highly similar to
HUMGA7A Human (clone GA733-2-2) carcinoma-associated antigen
GA733-2 mRNA.
ACCESSION AK026585
VERSION AK026585.1 GI:10439469
KEYWORDS Oligo clapping; fls (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III CDNA to
mRNA, clone_11b:KAT clone:KAT07515.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Watanabe, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,

TITLE Nakamura,Y., Isogai,T. and Sugano,S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1565)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..1565
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT07515"
/cell_line="KATO III"
/cell_type="signed-ring cell carcinoma"
/clone_lib="KAT"
/note="cloning vector pME18SFL3"
misc_feature
1..1565
/note="highly similar to HUMGA7A Human (clone GA733-2-2)
carcinoma-associated antigen GA733-2 mRNA"
BASE COUNT 468 a 321 c 367 g 409 t
ORIGIN
Query Match 66.9%; Score 449.8; DB 9; Length 1565;
Best Local Similarity 99.6%; Pred. No. 7.5e-65;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 156 TCAGATTAAGAGAGATGGTGCATAGGAGAACTCAATGCACTATATATATTGAA 215
1095 TGAGATTAAGAGAGATGGTGCATAGGAGAACTCAATGCACTATATATATTGAA 1154
QY 216 GATTATAGAAGAGAGAAATAGCAATTCACACAAATTCGCTGGAGAG 275
1155 GATTATAGAAGAGAGAAATAGCAATTCACACAAATTCGCTGGAGAG 1214
QY 276 AAGACATCTTTGAAGGCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 335
1215 AAGACATCTTTGAAGGCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 1274
QY 336 CCGTACTCAAAATATTAAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 395
1275 CCGTACTCAAAATATTAAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 1334
QY 396 AGTGTCTTATATGCAAGATCTAATGTAATCCAGAACTGGAGCTCATGTTAAAT 455
1335 AGTGTCTTATATGCAAGATCTAATGTAATCCAGAACTGGAGCTCATGTTAAAT 1394
QY 456 ATTATGTGTGAACATCAATGTCATTAATATGCTCCACAGTAAATGCGAAAA 515
1395 ATTATGTGTGAACATCAATGTCATTAATATGCTCCACAGTAAATGCGAAAA 1454
QY 516 CCGATTGTCATGTAAGAGCTGCTTCTATTACTGAGCTTGTACATACACTTTT 575
1455 CCGATTGTCATGTAAGAGCTGCTTCTATTACTGAGCTTGTACATACACTTTT 1514
QY 576 TATGACCTATGAATTAACACTTTTAACTGAA 608
1515 TATGACCTATGAATTAACACTTTTAACTGAA 1547
RESULT 7
HSKSA

LOCUS HSKSA 1504 bp mRNA linear PRI 22-MAR-1995
DEFINITION Human mRNA for adenocarcinoma-associated antigen (KSA).
ACCESSION X14758
VERSION X14758.1 GI:34186
KEYWORDS antigen; cell surface glycoprotein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1504)
Strnad,J., Hamilton,A.E., Beavers,L.S., Gamboa,G.C., Apelgren,L.D.,
Taber,L.D., Sportsman,J.R., Bunzl,T.F., Sharp,J.D. and Gadski,R.A.,
Molecular cloning and characterization of a human
adenocarcinoma/epithelial cell surface antigen complementary DNA
Cancer Res. 49 (2), 314-317 (1989)
JOURNAL MEDLINE 89089570
PUBMED 2463074
FEATURES
source Location/Qualifiers
1..1504
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="UCLA-P3"
/clone_lib="lambda gtl1."
155..1099
CDS
/note="KSA preproantigen peptide"
/codon_start=1
/protein_id="CAA32870.1"
/db_xref="SWISS-PROT:P16422"
/db_xref="GI:34187"
/translation="MAPPVYLAFGLLAAATATVAAOEECYENKLAIVGFEVNNR
OCQCTSVGQNTVCSKLAQKLVKAEKNSKLRKRPBGLDNBDLDPDDES
GFKAKQCGSTWCNVNAGVRRTDKDEIYCSERVRVWIIELKHKARPKYDSK
SLRTALQKEITRYOLDPKFTISILYENNVAIIRLVNSSQKTDVDAIVAYFEK
DYKGEISFKSKMDLVNGEOLDPGLTLYIVDEKARPEFSMGLKAGVLAIVVVV
MAVVGIVLVYISRRKKRYKAEIKEMGENHKLNA"
155..223
sig_peptide
/note="putative"
398..1096
mat_peptide
/product="mature KSA antigen (AA 59-291)"
224..1096
misc_feature
/note="KSA proantigen (AA 1-291)"
1504
polya_site
/note="polya site"
BASE COUNT 442 a 302 c 356 g 404 t
ORIGIN
Query Match 66.9%; Score 449.4; DB 9; Length 1504;
Best Local Similarity 99.8%; Pred. No. 8.8e-65;
Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 156 TCAGATTAAGAGAGATGGTGCATAGGAGAACTCAATGCACTATATATTGAA 215
1054 TGAGATTAAGAGAGATGGTGCATAGGAGAACTCAATGCACTATATATTGAA 1113
QY 216 GATTATAGAAGAGAGAAATAGCAATTCACACAAATTCGCTGGAGAG 275
1114 GATTATAGAAGAGAGAAATAGCAATTCACACAAATTCGCTGGAGAG 1173
QY 276 AAGACATCTTTGAAGGCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 335
1174 AAGACATCTTTGAAGGCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 1233
QY 1174 AAGACATCTTTGAAGGCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 1233
QY 336 CCGTACTCAAAATATTAAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 395
1234 CCGTACTCAAAATATTAAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 1293
QY 396 AGTGTCTTATATGCAAGATCTAATGTAATCCAGAACTGGAGCTCATGTTAAAT 455
1294 AGTGTCTTATATGCAAGATCTAATGTAATCCAGAACTGGAGCTCATGTTAAAT 1353
QY 456 ATTATGTGTGAACATCAATGTCATTAATATGCTCCACAGTAAATGCGAAAA 515
1354 ATTATGTGTGAACATCAATGTCATTAATATGCTCCACAGTAAATGCGAAAA 1413
DB

```

Oy 516 CTGATTTCGATGGAAGCGCTTTCATTTACTTGAGTCTGTGACATACACTTTT 575
Db 1414 CTGATTTCGATGGAAGCGCTTTCATTTACTTGAGTCTGTGACATACACTTTT 1473
Oy 576 TATGAGCTATGAAATTAACATTTTAACTG 606
Db 1474 TATGAGCTATGAAATTAACATTTTAACTG 1504

RESULT 8
HUMKSA 1504 bp mRNA linear PRI 11-JUN-1993
LOCUS Human adenocarcinoma-associated antigen (KSA) mRNA, complete cds.
DEFINITION M32325.1 GI:186775
ACCESSION M32325
VERSION M32325.1
KEYWORDS adenocarcinoma-associated antigen.
SOURCE Human cell line UCLA-P3, cDNA to mRNA, clone AG[1,1338,933].
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1504)
AUTHORS Strnad,J., Hamilton,A.E., Beavers,L.S., Gamboa,G.C., Apelgren,L.D.,
Taber,L.D., Sportsman,J.R., Bumol,T.F., Sharp,J.D. and Gadski,R.A.
TITLE Molecular cloning and characterization of a human
adenocarcinoma/epithelial cell surface antigen complementary DNA
JOURNAL Cancer Res. 49 (2), 314-317 (1989)
MEDLINE 89089570
PUBMED 2463074
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by J.D. Sharp, 22-FEB-1990.
There are a few base differences between the sequence presented
here and that which appears in entry M26481. The difference occurs
mostly in the Poly-A signal.
FEATURES
Source
Location/Qualifiers
1..1504
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..1504
/product="adenocarcinoma-associated antigen mRNA"
155..1099
/feature="adenocarcinoma-associated antigen precursor (KSA)"
/note="adenocarcinoma-associated antigen precursor (KSA)"
/codon_start=1
/protein_id="AA36151.1"
/db_xref="GI:307091"
/translation="MAPPYLAFGLIAAATATPAAOEVCNRYLAIVCPYNNR
OCQCTSVGAQNTVICSILAKCLVMAKEMNGSKGRARKEGALQNDGLYDPCDS
GLFRKAKCNGSTICWCYNTAGVTRTDKDTETCSERYRTWIIELHKAKEKYDSK
SLRTALOKETITTRQIDLPKFTSTLYENNVITIDLVONSSOKTQNDVDIADVAAYFEK
DVKGESLFHKKMDLVNGQDLDPQOTLIYVDEKAPESMGLAKAGYIAIVVVV
MAYVAGIVLVISIKKRMATYERKEIREKEMHRELNA"
155..223
/note="adenocarcinoma-associated antigen signal peptide"
mat_peptide 398..1096
/product="adenocarcinoma-associated antigen"
polya_signal 1486..1491
BASE COUNT 442 a 302 c 356 g 404 t
ORIGIN
Query Match 66.9%; Score 449.4; DB 9; Length 1504;
Best Local Similarity 99.8%; Pred. No. 8.8e-65;
Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 156 TCACATAAAGAGAGATGGTGCATAGGAGCACTCAATACATATTAATTGAA 215
Db 1054 TCACATAAAGAGAGATGGTGCATAGGAGCACTCAATACATATTAATTGAA 1113
Oy 216 GATTATAGAGAGAGAGAGATAGCAATATGACACAAATTAAGTGTGCTGCGAGCG 275
Db 1114 GATTATAGAGAGAGAGAGATAGCAATATGACACAAATTAAGTGTGCTGCGAGCG 1173
Oy 276 AAGCATCTTTGAAGGTCATGATTTGTTAATTAACTATATTTGTAATAGTAAA 335
Db 1174 AAGCATCTTTGAAGGTCATGATTTGTTAATTAACTATATTTGTAATAGTAAA 1233
Oy 336 CCTGTACTCAAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACACA 395
Db 1234 CCTGTACTCAAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACACA 1293
Oy 396 AGTGCTTATATATATGACATCTAATGTAAATCCGAACCTTGACATCTTAAATTT 455
Db 1294 AGTGCTTATATATATGACATCTAATGTAAATCCGAACCTTGACATCTTAAATTT 1353
Oy 456 ATTATGTATAACATTCATTAATGTGATTAATATGCTCCACAGTAATATCGAAAA 515
Db 1354 ATTATGTATAACATTCATTAATGTGATTAATATGCTCCACAGTAATATCGAAAA 1413
Oy 516 CTGATTTCGATGGAAGCGCTTTCATTTACTTGAGTCTGTGACATACACTTTT 575
Db 1414 CTGATTTCGATGGAAGCGCTTTCATTTACTTGAGTCTGTGACATACACTTTT 1473
Oy 576 TATGAGCTATGAAATTAACATTTTAACTG 606
Db 1474 TATGAGCTATGAAATTAACATTTTAACTG 1504

RESULT 9
HUMGA7A 1457 bp mRNA linear PRI 27-APR-1993
LOCUS Human (clone GA733-2-2) carcinoma-associated antigen GA733-2 mRNA,
complete cds.
ACCESSION M33011
VERSION M33011.1 GI:182895
KEYWORDS carcinoma-associated antigen GA733-2.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1457)
AUTHORS Szala,S., Froehlich,M., Scollon,M., Kasal,Y., Steplewski,L.,
Koprowski,H. and Linenbach,A.J.
TITLE Molecular cloning of cDNA for the carcinoma-associated antigen
GA733-2
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (9), 3542-3546 (1990)
MEDLINE 90239051
PUBMED 2333300
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by A.J. Linenbach, 22-MAR-1990.
FEATURES
Source
Location/Qualifiers
1..1457
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="SW948"
/cell_type="colorectal carcinoma"
<1..1457
mat_peptide 108..155
/note="putative"
polya_signal 1439..1444
BASE COUNT 437 a 285 c 339 g 396 t
ORIGIN
Query Match 66.6%; Score 447.8; DB 9; Length 1457;
Best Local Similarity 99.6%; Pred. No. 1.6e-64;

```

Matches	449:	Conservative	0:	Mismatches	2:	Indels	0:	Gaps	0:
QY	156	TCAGATAAAGGAGATGGGTGAGATGCATAGGCACTCAATGCATATATATATTGAA	215						
Db	1007	TGAGATTAAGGAGATGGGTGAGATGCATAGGGAACCTCAATGCATATATATATTGAA	1066						
QY	216	GATTATAGAGAGGGAATATGCAATGGACCAATATACAAATGTGTGGTGGGAGC	275						
Db	1067	GATTATAGAGAGGGAATATGCAATGGACCAATATACAAATGTGTGGTGGGAGC	1126						
QY	276	AAGACATCTTTGAAGGTCATGAGTTTGTACTTTAAATCATATATATTTGTAACTGAAA	335						
Db	1127	AAGACATCTTTGAAGGTCATGAGTTTGTACTTTAAATCATATATTTGTAACTGAAA	1186						
QY	336	CCTGATACCAAAATATTAAGCACTTGAACCTGTTTACCAATCTTGAATTTGGACCA	395						
Db	1187	CCTGATACCAAAATATTAAGCACTTGAACCTGTTTACCAATCTTGAATTTGGACCA	1246						
QY	396	AGTGTCTTATATATGCAGATCTTAATGTAAAAATCCAGAACTTGGACTGCATCTTAAATT	455						
Db	1247	AGTGTCTTATATATGCAGATCTTAATGTAAAAATCCAGAACTTGGACTGCATCTTAAATT	1306						
QY	456	ATTTATGTGTAACTTCAAAATGTGTGCATTTAAATATGCTCCACAGTAAAAATCTGAAAA	515						
Db	1307	ATTTATGTGTAACTTCAAAATGTGTGCATTTAAATATGCTCCACAGTAAAAATCTGAAAA	1366						
QY	516	CTGATTTTGTGATTTGAAGAGCTGCTTTCATATTTACATTTGAGTCTTTGATACATACACTTTT	575						
Db	1367	CTGATTTTGTGATTTGAAGAGCTGCTTTCATATTTACATTTGAGTCTTTGATACATACACTTTT	1426						
QY	576	TATGAGCTATGAAATATAAATCTTTAAACTG	606						
Db	1427	TATGAGCTATGAAATATAAATCTTTAAACTG	1457						

LOCUS	AX395799/c	637 bp	DNA	linear	PAT 18-MAY-2002
DEFINITION	Sequence	14	from Patent WO0212328.		
ACCESSION	AX395799				
VERSION	AX395799.1	GI:21066546			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	King, G.E., Meagher, M.J., Xu, J. and Secrist, H.				
TITLE	Compositions and methods for the therapy and diagnosis of colon cancer				
JOURNAL	Patent; WO 0212328-A 14 14-FEB-2002;				
FEATURES	CORIXA CORPORATION (US)				
source	Location/Qualifiers				
	1..637				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	205 a 136 c 84 g 212 t				
ORIGIN					
Query Match	66.0%;	Score 443.4;	DB 6;	Length 637;	
Best Local Similarity	99.8%;	Pred. NO. 1e-63;			
Matches 444;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	156 TCAGATTAAGGAGATGCGTGAGATGCATATGGGAACTCATGCATATCTATTAATTGGAA	215			
Db	445 TAGAGTTAAAGGAGATGGGGAGATGCTATGGGAACTCATGCATATCTATTAATTGGAA	386			
QY	216 GATTATAGAGAAGGGAATTACCAATGACACCAAAATTCAATATGCTGCGTGGAGCG	275			
Db	385 GATTATAGAGAAGGGAATTACCAATGACACCAAAATTCAATATGCTGCTGGAGCG	326			
QY	276 AAGACATCTTTGAGGCTCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA	335			

Db	325	AASACATCTTTGGAAGGTCATGAGTTTGTGTAGTTTAACATCATATATTTGTAAATAGTGA	266
QY	336	CCTGTACTCAAAATATTAAGCAGCGTTGAACTGCGTTTACCAATCTTGAATTTTGACCACA	395
Db	265	CCTGTACTCAAAATATTAAGCAGCGTTGAAACGCGTTTACCAATCTTGAATTTTGACCACA	206
QY	396	AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTTGCACTCGTTTAAATTT	455
Db	205	AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTTGCACTCGTTTAAATTT	146
QY	456	ATTATATGTGTAACTTCAATGTGTGCATTAAATATGCTTCCACAGTAAATCTGAAAA	515
Db	145	ATTATATGTGTAACTTCAAAATGTGTGATTAAATATGCTTCCACAGTAAATCTGAAAA	86
QY	516	CTGATTTTGTGATGAAACGTCGCTTTCATTTACTTGTAGTCTGTACATACATACTTTTT	575
Db	85	CTGATTTTGTGATGAAACGTCGCTTTCATTTACTTGTAGTCTGTACATACATACTTTTT	26
QY	576	TATGAGCTATGAAATTAACAATTTT	600
Db	25	TATGAGCTATGAAATTAACAATTTT	1

LOCUS	AX397344/C	481 bp	DNA	linear	PAT 18-MAY-2002
DEFINITION	Sequence 1559 from Patent WO0212328.				
ACCESSION	AX397344				
VERSION	AX397344.1	GI:21068091			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 King, G.E., Meagher, M.J., Xu, J. and Secrist, H.				
TITLE	Compositions and methods for the therapy and diagnosis of colon cancer				
JOURNAL	Patent: WO 0212328-A 1559 14-FEB-2002;				
FEATURES	CORIXA CORPORATION (US)				
source	Location/Qualifiers				
	1..481				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	149 a 93 c 64 g 174 t	1	others		
ORIGIN					
Query Match	55 88: Score 442.4: DB 6: Length 481:				
Best Local Similarity	99.68: Pred. NO. 1.6e-63:				
Matches 443: Conservative	0: Mismatches 2: Indels 0: Gaps 0:				
QY 156	TCAGATAAAGAGATGGGTGAGATGCATAGGGAAGCTCAATGCATACATATATATTAATTGAA	215			
DB 445	TCAGATAAAGAGATGGGTGAGATGCATAGGGAAGCTCAATGCATATATATTAATTGAA	386			
QY 216	GATTATAGAGAAGGAATTAGCAATAGCACCAATTCATCATGTGTGGTGGGAGC	275			
DB 385	GATTATAGAGAAGGAATTAGCAATAGCACCAATTCATCATGTGTGGTGGGAGC	326			
QY 276	AAGACATCTTTAAGGTCATGAGTTTGTAGTTTAACATCATATATTTGTAAATAGTAAA	335			
DB 325	AAGACATCTTTGAGGTCATGAGTTTGTAGTTTAACATCATATATTTGTAAATAGTAAA	266			
QY 336	CCTGTACTAAAAATATAGACAGCTTGAACCTGGCTTTACCATCTTTGAAATTTGACCCCA	395			
DB 265	CCTGTACTAAAAATATAGACAGCTTGAACCTGGCTTTACCATCTTTGAAATTTGACCCCA	206			
QY 396	AGTGTCTTATATATGAGATCTAATGTAAATCCAGAACTGGAGCTCCATCGTTAAAT	455			
DB 205	AGTGTCTTATATATGAGATCTAATGTAAATCCAGAACTGGAGCTCCATCGTTAAAT	146			
QY 456	ATTATGTGTAAACATTCAAATGTGTGCATTTAAATATGCTTCCACAGTAAATCTGAAAAA	515			

Db 145 ATTATGTGTAACATCAATGATGTCATTAATATGCTCCACAGTAAATCTGAAAA 86
 Oy 516 CTGATTTGTGATGAAGCTGCTTCTATTTACTGATGCTGTGACATACACTTTT 575
 Db 85 CTGATTTGTGATGAAGCTGCTTCTATTTACTGATGCTGTGACATACACTTTT 26
 Oy 576 TATGAGCTATGAATTAACATTTT 600
 Db 25 TATGAGCTATGAATTAACATTTT 1

RESULT 12
 AX198887/c AX198887 545 bp DNA linear PAT 29-AUG-2001
 LOCUS Sequence 1342 from Patent WO0151513.
 DEFINITION AX198887
 ACCESSION AX198887
 VERSION AX198887.1 GI:15389213
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 545)
 AUTHORS Algate,P.A.
 TITLE Ovarian tumor-associated sequences
 JOURNAL Patent: WO 0151513-A 1342 19-JUL-2001;
 CORIXA CORPORATION (US)
 FEATURES
 source location/Qualifiers
 1..545
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 163 a 112 c 89 g 177 t 4 others
 ORIGIN

Query Match 65.8%; Score 442.2; DB 6; Length 545;
 Best Local Similarity 96.0%; Pred. No. 1.7e-63;
 Matches 453; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Oy 129 GAATTTTTCGTCGCTTTTCCGTTCCAGTAAAGAGATGGGTGATGATGAGGA 188
 Db 472 GGAATTCGCCNTTACCGTGGTGGCGCGAGGTAAAGAGATGGGTGATGATGAGGA 413
 Oy 189 ACTCAATGCATTAATTAATTTGAAGATTATAGAGAAGGAATAGCAATGGACAC 248
 Db 412 ACTCAATGCATTAATTAATTTGAAGATTATAGAGAAGGAATAGCAATGGACAC 353
 Oy 249 AAATTACAATGTGTGCGGTGGAGCAAGACATCTTTGAAGGTGATGATTTT 308
 Db 352 AAATTACAATGTGTGCGGTGGAGCAAGACATCTTTGAAGGTGATGATTTT 293
 Oy 309 TAACATCATATTTTGTATATAGTGAACCTGTACTCAAAATTAAGCAGCTTGAAC 368
 Db 292 TAACATCATATTTTGTATATAGTGAACCTGTACTCAAAATTAAGCAGCTTGAAC 233
 Oy 369 CTTTACCATCTTGAATTTTGGACACAAGTGTCTATATATGACATCTAATTAAT 428
 Db 232 CTTTACCATCTTGAATTTTGGACACAAGTGTCTATATATGACATCTAATTAAT 173
 Oy 429 CAGAAGCTGACCTCATGTTAAATTTATTTATGTGAACATTCAAATGTGCAATTA 488
 Db 172 CAGAAGCTGACCTCATGTTAAATTTATTTATGTGAACATTCAAATGTGCAATTA 113
 Oy 489 TATGCTTGCACAGTAAATCTGAAAAAGTATTTGTGATGAAGCTGCTTTCTATTTA 548
 Db 112 TATGCTTGCACAGTAAATCTGAAAAAGTATTTGTGATGAAGCTGCTTTCTATTTA 53
 Oy 549 CTTGAGCTTGTGACATACACTTTTATGAGCTATGAAAAATTAACATTTT 600
 Db 52 CTTGAGCTTGTGACATACACTTTTATGAGCTATGAAAAATTAACATTTT 1

RESULT 13
 HUMMAKSL14

LOCUS HUMMAKSL14 1511 bp mRNA linear PRI 09-JAN-1995
 DEFINITION Homo sapiens KSL/4 antigen mRNA, complete cds.
 ACCESSION M26481
 VERSION M26481.1 GI:619789
 KEYWORDS KSL/4 antigen.
 SOURCE Homo sapiens lung adenocarcinoma CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1511)
 AUTHORS Perez,M.S. and Walker,L.E.
 TITLE Isolation and characterization of a cDNA encoding the KSL/4
 JOURNAL epithelial carcinoma marker
 MEDLINE J. Immunol. 142 (10), 3662-3667 (1989)
 PUBMED 89235177
 COMMENT 2469722
 On Jan 10, 1995 this sequence version replaced gi:341508.
 FEATURES
 source location/Qualifiers
 1..1511
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="UCLA-P3"
 /cell_type="B lymphoblast"
 /tissue_type="lung adenocarcinoma"
 160..1104
 /codon_start=1
 /product="KSL/4 antigen"
 /protein_id="AA59543.1"
 /db_xref="GI:619790"
 /translation="MAPQVLAIFGLLAATATATVAAGCEVCENYKLVNCPVNNR
 QOCCTVGAQNTVICSKLAAKLVKAMNKSILGRKPGALONNDGLDPDDES
 GLFRKQCNSTWCNVTAGVRDKDTEITGSESVATYIIEIKKAREKPYDSK
 SLRTALQKEITTRQIDPKFTITSLYENNVITIDVONSOKTQNDVDIADVAVFEEK
 DYKSESLFHSKKMLDTVNGEDLDIPOTLLIYYDERKPEESMCKRGVLAIVVVV
 MAVVAGIVLVISRRKRMAKTEKRIKEMHHEHLNA"
 sig_peptide 160..228
 mat_peptide 229..1101
 polyA_signal 1490..1495
 /note="putative"
 /product="KSL/4 antigen"
 /note="putative"
 BASE COUNT 443 a 305 c 359 g 404 t

Query Match 65.4%; Score 439.4; DB 9; Length 1511;
 Best Local Similarity 99.6%; Pred. No. 3.9e-63;
 Matches 451; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 156 TCAGATTAAGGAGATGGTGAATGCATGATGGAACATGATGATGATGATGATGATG 215
 Db 1059 TCAGATTAAGGAGATGGTGAATGCATGATGGAACATGATGATGATGATGATG 1118
 Oy 216 GATTATAGAAGAGGAATAGCAAAATGAGACAAATTAACAATCTGTGCGTGGAGC 275
 Db 1119 GATTATAGAAGAGGAATAGCAAAATGAGACAAATTAACAATCTGTGCGTGGAGC 1178
 Oy 276 AAGACATCTTTAAGTCTAGAGTTTGTATTTAATCATCATATATTTTATAGTGA 335
 Db 1179 AAGACATCTTTAAGTCTAGAGTTTGTATTTAATCATCATATATTTTATAGTGA 1238
 Oy 336 CCTGTACTCAAAATATATAGCAGTTGAACCTGCTTACCAATCTTGAATTTGACACA 395
 Db 1239 CCTGTACTCAAAATATATAGCAGTTGAACCTGCTTACCAATCTTGAATTTGACACA 1298
 Oy 396 AGTGTCTTATATATGAGATCTAATGTAAATCCAGAACCTTGACATCTGTTAAAT 455
 Db 1299 AGTGTCTTATATATGAGATCTAATGTAAATCCAGAACCTTGACATCTGTTAAAT 1358
 Oy 456 ATTATGTGTAACATTTCAATGTGTGATTAATATATGCTTCCACAGTAAATCTGAAAA 515
 Db 1359 ATTATGTGTAACATTTCAATGTGTGATTAATATATGCTTCCACAGTAAATCTGAAAA 1418
 Oy 516 CTGATTTGTGATGAAGCTGCTTCTATTTACTGATGATGATGATGATGATGATG 575

```

Db 1419 CTCATTGTCATTGAAGCGCTTCTATTACTGAGCTTGTACATACATACATTTT 1478
QY 576 TATGACCTATGAATAAATCAATTTTAAACTGAA 608
Db 1479 TATGACCTATG-AATTAACATTTTAACTGAA 1510

RESULT 14
AX209414/C 439 bp DNA linear PAT 31-AUG-2001
LOCUS AX209414.1
DEFINITION Sequence 1254 from Patent WO0157207.
ACCESSION AX209414
VERSION AX209414.1 GI:15423837
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian
cancer
JOURNAL Patent: WO 0157207-A 1254 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. 439
Location/Qualifiers
BASE COUNT 142 a 80 c 60 g 157 t
ORIGIN
Query Match 65.3%; Score 439; DB 6; Length 439;
Best Local Similarity 100.0%; Pred. No. 6e-63;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AAGAGATGCGGTGAGATGATAGGAGACTGATGATATATATATTTGAAGATTAT 221
Db 439 AAGAGATGCGGTGAGATGATAGGAGACTGATGATATATATATTTGAAGATTAT 380
QY 222 AGAGAGAGGAATATGACAAATGACAAATTTCAATATGCTGTGGGTGGAGAGACA 281
Db 379 AGAGAGAGGAATATGACAAATGACAAATTTCAATATGCTGTGGGTGGAGAGACA 320
QY 282 TCTTTGAGGTCATGAGATTGTAGTTAAATCATATATTTGTAATAGGAACCTGTA 341
Db 319 TCTTTGAGGTCATGAGATTGTAGTTAAATCATATATTTGTAATAGGAACCTGTA 260
QY 342 CTCAAATATTAAGCAGCTTGAACCTGTTACCAATCTTGAATTTGACCAAGTGTG 401
Db 259 CTCAAATATTAAGCAGCTTGAACCTGTTACCAATCTTGAATTTGACCAAGTGTG 200
QY 402 TTATATATGAGATCTAATGTAATCCGAATCTGAGCTCCATCGTTAAATTTATTTAT 461
Db 199 TTATATATGAGATCTAATGTAATCCGAATCTGAGCTCCATCGTTAAATTTATTTAT 140
QY 462 GGTCAACATTTCAAAATGTCATTAATATGCTTCCACAGTAATAAATCTGAAAACTGATT 521
Db 139 GGTCAACATTTCAAAATGTCATTAATATGCTTCCACAGTAATAAATCTGAAAACTGATT 80
QY 522 TGTGATGTAAGAGCTGCTTTCTATTACTTGAAGCTTGTGATACATACATCTTTTATGAG 561
Db 79 TGTGATGTAAGAGCTGCTTTCTATTACTTGAAGCTTGTGATACATACATCTTTTATGAG 20
QY 582 CTATGAATATAAATCAATTTT 600
Db 19 CTATGAATATAAATCAATTTT 1

RESULT 15
HUMCSAE 1372 bp mRNA linear PRI 27-APR-1993
LOCUS HUMCSAE
DEFINITION Human epithelial glycoprotein (EGP) mRNA, complete cds.

```

```

ACCESSION M32306
VERSION M32306.1 GI:181132
KEYWORDS cell surface antigen; cell surface glycoprotein; epithelial
glycoprotein; transmembrane protein; tumor-associated antigen.
SOURCE Human colon carcinoma cell line HT-29, cDNA to mRNA, (library
HT-29/cdm8).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1372)
AUTHORS Simon, B., Podolsky, D. K., Moldenhauer, G., Isselbacher, K. J.,
Galloni, C. E., S. and Brand, S. J.
TITLE Epithelial glycoprotein is a member of a family of epithelial cell
surface antigens homologous to nidogen, a matrix adhesion protein
Proc. Natl. Acad. Sci. U.S.A. 87 (7), 2755-2759 (1990)
JOURNAL 90207276
MEDLINE 2108441
PUBMED
COMMENT Draft entry and computer-readable sequence (1) kindly submitted by
B. Simon, 23-FEB-1990.
FEATURES
source 1. 1372
Location/Qualifiers
CDS mRNA
1. 1372
/db_xref="taxon:9606"
<1. 1372
/product="EGP mRNA"
22. 966
/feature="epithelial glycoprotein (EGP) precursor"
/note="epithelial glycoprotein signal peptide"
mat_peptide 85. 963
BASE COUNT 436 a 233 c 317 g 386 t
ORIGIN
Query Match 65.2%; Score 438.4; DB 5; Length 1372;
Best Local Similarity 98.7%; Pred. No. 5.8e-63;
Matches 442; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 156 TCAGATTAAGAGATGGGTGAGATGATAGGAGACTCAATGATCAATATATATTTGAA 215
Db 921 TGAGATTAAGAGATGGGTGAGATGATAGGAGACTCAATGATCAATATATATTTGAA 980
QY 216 GATTATGAAGAAGGAATATGACAAATTTCAATATGCTGTGGGTGGAGC 275
Db 981 GATTATGAAGAAGGAATATGACAAATTTCAATATGCTGTGGGTGGAGC 1040
QY 276 AAGACATCTTTGAAGTTCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 335
Db 1041 AAGACATCTTTGAAGTTCATGAGTTGTTAGTTAAATCATATATTTGTAATAGTGA 1100
QY 336 CCGTACTCAAAATATTAAGCAGCTTGAACCTGCTTTACCAATCTTGAATTTGACAC 395
Db 1101 CCGTACTCAAAATATTAAGCAGCTTGAACCTGCTTTACCAATCTTGAATTTGACAC 1160
QY 396 AGTGTCTTATATATGAGATCTAATGTAATCCGAATCTGAGCTCCATCGTTAAATTT 455
Db 1161 AGTGTCTTATATATGAGATCTAATGTAATCCGAATCTGAGCTCCATCGTTAAATTT 1220
QY 456 ATTATGTGTAACATTTCAATGTCATTAATATGCTTCCACAGTAATAAATCTGAAAA 515
Db 1221 ATTATGTGTAACATTTCAATGTCATTAATATGCTTCCACAGTAATAAATCTGAAAA 1280
QY 516 CTGATTTGATGTAAGAGCTGCTTTCTATTACTTGAAGCTTGTGATACATACATCTTTT 575

```


Db 1281 CTGATTGTGATTGAAGCTGCTTCTATTACTTGAGCTCTGTACATACATCTTTT 1340
Oy 576 TATGAGCTATGAATATAACATTTTAA 603
|||||
Db 1341 TATGAGCTATGAATATAACATTTTAA 1368
|||||

Search completed: November 26, 2002, 22:48:44
Job time : 1288.24 secs

THIS PAGE BLANK (USPTO)

PT expressed in normal and malignant endometrium -
XX
PS Claim 1, page 97, 233pp, English.

CC The invention relates to diagnosing endometrial cancer in a subject
CC suspected of having endometrial cancer comprising determining the
CC expression of a set of nucleic acid molecules or expression products in
CC an endometrial sample suspected of being cancerous, where the set of
CC nucleic acid molecules comprises at least 2 nucleic acid molecules
CC selected from 50 fully defined sequences as given in the specification.
CC The nucleic acids are used as an array of at least 2 of the 50
CC nucleic acids bound to a solid substrate. Also included is a solid-phase
CC protein microarray comprising at least 2 antibodies or its antigen
CC binding fragments, that specifically bind at least 2 different
CC polypeptides from the 50 fully defined sequences as given in the
CC specification, fixed to a solid substrate. The methods and arrays are
CC useful for the diagnosis of endometrial cancer, selecting and monitoring
CC treatment regimes and identification of lead compounds useful for the
CC treatment of endometrial cancer. The present sequence is one of 50
CC genes differentially expressed between cancerous and non-cancerous
CC samples.

XX Sequence 672 BP; 222 A; 85 C; 100 G; 265 T; 0 other;

Query Match 100.0%; Score 672; DB 24; Length 672;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATATACCTTTCTTTTATTTAATTAAGTCTTAATAAGTGGGAAAAATATCTT 60
DB .1 AATATACCTTTCTTTTATTTAATAAGCTTAATAAGTGGGAAAAATATCTT 60
QY 61 GGTGTCCTTAATTTTATTTAATTAAGTCTTAATAAGTGGGAAAAATATCTT 120
DB 61 GGTGTCCTTAATTTTATTTAATTAAGTCTTAATAAGTGGGAAAAATATCTT 120
QY 121 ATTATTAAGATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTT 180
DB 121 ATTATTAAGATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTT 180
QY 181 CATAGGGAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 240
DB 181 CATAGGGAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 240
QY 241 ATGGAACAATTAACAATTAACAATTAACAATTAACAATTAACAATTAACAATTAACA 300
DB 241 ATGGAACAATTAACAATTAACAATTAACAATTAACAATTAACAATTAACAATTAACA 300
QY 301 TGTATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 360
DB 301 TGTATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 360
QY 361 GAAGTGGCTTACCAATCTTGAATTTGACCAAGTGTCTTATATGACATCAATCAAT 420
DB 361 GAAGTGGCTTACCAATCTTGAATTTGACCAAGTGTCTTATATGACATCAATCAAT 420
QY 421 GTAAATCCAGACTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 480
DB 421 GTAAATCCAGACTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 480
QY 481 GCATTAATATGCTTCACAGTAATAATCTGAATAATCTGAATAATCTGAATAATCTGCTT 540
DB 481 GCATTAATATGCTTCACAGTAATAATCTGAATAATCTGAATAATCTGAATAATCTGCTT 540
QY 541 TCTATTTACTGAGCTTGTGACATCAATCTTTTATGAGCTATGAATAAACAATTTT 600
DB 541 TCTATTTACTGAGCTTGTGACATCAATCTTTTATGAGCTATGAATAAACAATTTT 600
QY 601 AAACGTAATTTCTTAATTTGACATTTCAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
DB 601 AAACGTAATTTCTTAATTTGACATTTCAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
QY 661 TTTTCTTTGAGA 672

DB 661 TTTTCTTTGAGA 672

RESULT 2

AA524010/C
ID AA524010 standard; cDNA; 621 BP.

AA524010;

07-NOV-2001 (first entry)

Human ovarian PCR-subtracted cDNA library clone #191.

Immunogenic protein: cancer; ovarian tumour; T-cell stimulation; ss;
gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;

primer; probe.

Homo sapiens.

MO200157207-A2.

09-AUG-2001.

05-FEB-2001; 2001MO-US03733.

04-FEB-2000; 2000US-0180403.

28-MAR-2000; 2000US-0192745.

(CORI-) CORIXA CORP.

Algate PA, Mannion J;

WPI; 2001-488879/53.

New polynucleotides encoding ovarian tumour proteins, useful for
treating ovarian cancer, and as probes, primers, and markers of cancer

progression

Example 1: page 140-141; 378pp; English.

XX The invention comprises compositions used for the therapy and diagnosis
XX of ovarian cancer. The compositions comprise one or more ovarian tumour
XX proteins, their associated polynucleotides, or immunogenic portions of
XX the proteins. The ovarian tumour polynucleotides and polypeptides are
XX useful for stimulating and/or expanding T cells specific for a tumour
XX protein. They are also useful for inhibiting the development of cancer in
XX a patient with an ovarian tumour DNA or protein by incubating isolated
XX T-cells allowing them to proliferate, and administering to the patient.
XX The sequences can be used as markers for cancer, for example, to monitor
XX ovarian cancer progression. Probes and primers are useful in nucleic acid
XX hybridisation, in detecting the presence of complementary sequences in a
XX given sample, for preparing mutant species and for preparing other
XX genetic constructions. Sequences AA523820-AA525231 and AA525328-AA525549
XX represent human ovarian tumour protein cDNA clones.

Sequence 621 BP; 207 A; 118 C; 82 G; 195 T; 19 other;

Query Match 68.7%; Score 461.6; DB 22; Length 621;
Best Local Similarity 91.3%; Pred. No. 2.8e-71;
Matches 525; Conservative 0; Mismatches 44; Indels 6; Gaps 4;

QY 26 AATTAAGTCTTATTAATGCGAAAAAATATGCTGCTTAAATTTTCAATTTTAT 85
DB 26 AATTAAGTCTTATTAATGCGAAAAAATTTTCTGTG---TTCTTTAATTTTCAATTTTAT 85
QY 86 TAATTAATTTTCAATGAGCAAAAGATTAATTAATTTTCTGTGCT 145
DB 86 TAATTAATTTTCAATGAGCAAAAGATTAATTAATTTTCTGTGCT 145
QY 512 TAANGCATTTTTCANAGATG-ACNAAAGATTTGAATAATATATNNG-AVTTTCTGTGCT 455
DB 512 TAANGCATTTTTCANAGATG-ACNAAAGATTTGAATAATATATNNG-AVTTTCTGTGCT 455
QY 146 TTTTCTGTTTCAGATTAAGGAGATGGGTGAGATGCATAGGGAATCAATCAATCAAT 205
DB 146 TTTTCTGTTTCAGATTAAGGAGATGGGTGAGATGCATAGGGAATCAATCAATCAAT 205
QY 454 TTTTCCNGTTTCAGATTAAGGAGATGGGTGAGATGCATAGGGAATCAATCAATCAAT 396
DB 454 TTTTCCNGTTTCAGATTAAGGAGATGGGTGAGATGCATAGGGAATCAATCAATCAAT 396

Oy	206	ATPTTTGAAGATTATGGAAGGAAGGAATPAGCAATATGACACCAAAATTACAAATGCTGCT	265
Dd	335	ATACCTTGAAGATTATGGAAGGAAGGAATPAGCAATATGACACCAAAATTACAAATGCTGCT	336
Oy	266	GCGTGGAGCAGACACATCTTTGAAGGCTCAGATGTTGTAGTTTAACATCATATATTTGT	325
Dd	335	GCGTGGAGCAGACACATCTTTGAAGGCTCAGATGTTGTAGTTTAACATCATATATTTGT	276
Oy	326	AATAGTGAACCTGTACTCAAAATATPAGCAGCTTGAACCTGGCTTTACCAATCTTGAA	385
Dd	275	AATAGTGAACCTGTACTCAAAATATPAGCAGCTTGAACCTGGCTTTACCAATCTTGAA	216
Oy	386	TTTACACACAGGCTCTTATATATGAGATCTPAGTAAATCCAGAACTGGACATCAT	445
Dd	215	TTTACACACAGGCTCTTATATATGAGATCTPAGTAAATCCAGAACTGGACATCAT	156
Oy	446	CGTTTAAATATTTATGATGTACATTTCAATGATGTGCATTAATATGCTTCCACAGTAA	505
Dd	155	CGTTTAAATATTTATGATGTACATTTCAATGATGTGCATTAATATGCTTCCACAGTAA	96
Oy	506	ATCGAAAAAAGTGAATTTGATGATGAAGAGCTGCTTCTTACTTGAAGTCTTGACATA	565
Dd	95	ATCGAAAAAAGTGAATTTGATGATGAAGAGCTGCTTCTTACTTGAAGTCTTGACATA	36
Oy	566	CATACCTTTTATGAGCTATGAATTAACATTTT	600
Dd	35	CATGCTTTTATGAGCTATGAATTAACATTTT	1
RESULT 3			
AAH82577/c			
ID	AAH82577	standard; cDNA; 621 bp.	
AC	AAH82577;		
XX	25-SEP-2001	(first entry)	
XX	Human ovarian tumour associated polynucleotide sequence SEQ ID NO:201.		
XX	Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;		
KW	immunogenic; vaccine; ss.		
XX	Homo sapiens.		
OS	WO200151513-A2.		
PN	19-JUL-2001.		
XX	16-JAN-2001; 2001WO-US01575.		
XX	14-JAN-2000; 2000US-0176722.		
XX	(CORI-) CORIXA CORP.		
PA	Algate PA;		
EI	WPI; 2001-425866/45.		
XX	Novel ovarian tumor proteins, and nucleic acids encoding them, used to		
PT	treat and diagnose cancers, particularly ovarian cancer -		
XX	Claim 5; Page 99; 338bp; English.		
XX	AAH82377 to AAH83878 represent human ovarian tumour-associated		
CC	polynucleotide sequences which encode ovarian tumour proteins. The		
CC	ovarian tumour protein and polynucleotide sequences have cytostatic		
CC	activity, and can be used in gene therapy and vaccine production. The		
CC	ovarian tumour proteins and polynucleotides can be used to inhibit		
CC	the development of cancer, particularly ovarian cancer. They can also		
CC	be used to diagnose the onset and progression of cancer.		
XX	Sequence 621 bp; 207 A; 118 C; 82 G; 195 T; 19 other;		

Query Match	68.7%	Score 461.6	DB 22	Length 621
Best Local Similarity	91.3%	Pred. No. 2,8e-71		
Matches 525	Conservative 0	Mismatches 44	Indels 6	Gaps
Oy	26	AAATAGGCTATTAATAGTGGGAAAAAATATCTGTGCTCCCTTAATTCATTTTAT	85	
Db	569	AAATAGGCTATTAATAGTGGGAAAAAATTTCTGTG---TTCTTTATTTCTTTTAT	513	
Oy	86	TAATACATTTTTCAGAAATGAACAAAAGATTGAAAAATTTATAGATTTTTTCTGTGCT	145	
Db	512	TAAGCATATTTTTCANGATG-ACNAAAGATTGAAAAATTTATNAG-ATTTTTCTGTGCT	455	
Oy	146	TTTTCTCTTTTCAGATTAAGGAGATGGGTGAGATGCATAGGAGCTCAATGCATACAT	205	
Db	454	TTTTCCGCTTTTCAGATTAAGGAGGGTGAGATGATG-GGAACTCAATGCATACAT	396	
Oy	206	ATAATTTGAAGATATATAGAAAGGCAAAATAGCAATAGACACAATTTCAATATGTGT	265	
Db	395	ATACCTTGAAGATATATAGAAAGGCAAAATAGCAATAGGCAAAATTTCAATATGTGT	336	
Oy	266	GCGTGGGACGACGACATCTTTGAAGTGCATGACTTTGTTAGTTTAAATCATATATTTGT	325	
Db	335	GCGGGGACGAGACATCTTTGAAGGTCATGACTTTGTTAGTTTAAATCATATATTTGT	276	
Oy	326	AATAGTGAACCTCTACTCAAAAATATTAAGCAGCTTGAACTGGCTTTACCAATCTTGAA	385	
Db	275	AATAGTGAACCTCTACTCAAAAATATTAAGCAGCTTGAACTGGCTTTTNCATCTTGAA	216	
Oy	386	TTTGAACCAAGTCTTATATATATGAGATCTATATGTAAATTCAGAACTTGGACTCAT	445	
Db	215	TTTGGCCCAAGTCTTATATATATGAGATCTATATGTAAATTCAGAACTTGGACTCAT	156	
Oy	446	CGTTAAATTTTATATGTGTGAACATTCAAATGTGCATTAATATATGCTTCACAGTAA	505	
Db	155	CGTTAAATTTTATATGTGTGAACATTCAAATGTGCATTAATATATGCTTCACAGTAA	96	
Oy	506	ATTCGAAAAACGTATTTGTATTTGAAAGCTGCGCTTTCTATTTACTTGAGCTTGTACATA	565	
Db	95	ATTCGAAAAACGTATTTGTATTTGAAAGCTGCGCTTTCTATTTACTTGAGCTTGTACATA	36	
Oy	566	CATACCTTTTATGAGCTATGAAATAAACATTTT	600	
Db	35	CATGCTTTTATGAGCTATGAAATAAACATTTT	1	
RESULT 4				
AAZ77543				
ID	AAZ77543 standard; cDNA; 1545 BP.			
XX	AAZ77543;			
XX	10-APR-2000 (first entry)			
XX	Human ovarian tumor cDNA library derived EST fragment 94.			
XX	Expressed sequence tag; EST; human; ovarian tumor; anticancer;			
XX	gene therapy; treatment; ss.			
XX	Homo sapiens.			
XX	DE19817557-AL.			
XX	21-OCT-1999.			
XX	09-APR-1998; 98DE-1017557.			
XX	09-APR-1998; 98DE-1017557.			
XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.			
XX	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;			
XX				

Db 1328 AGGCTCTTATATGACAGATCTAATGTAAATCCAGAACTTGACCTCCATGCTTAAATTT 1387
 QY 456 ATTATATGTAACTTCAATATGTGATTAATATGCTTCACAGTAATCTGAAAAA 515
 Db 1388 ATTTATGTGTACATTCATAATGTGTGATTAATATGCTTCACAGTAATCTGAAAAA 1447
 QY 516 CTGATTTGTGATTTGAAGCTGCTTTCTATTACTTGAGTCTTGTACATACATCTTTT 575
 Db 1448 CTGATTTGTGATTTGAAGCTGCTTTCTATTACTTGAGTCTTGTACATACATCTTTT 1507
 QY 576 TATGAGCTATGAAATTAACATTTTAAACTGAA 608
 Db 1508 TATGAGCTATGAAATTAACATTTTAAACTGAA 1540

RESULT 6

AAH33540
 ID AAH33540 standard; cDNA; 1597 BP.

AC AAH33540;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:596.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 4; ss.

OS Homo sapiens.

XX MO200122920-A2.

XX PD 05-APR-2001.

PE 28-SEP-2000; 2000MO-US26524.

XX PR 29-SEP-1999; 99US-0157137.

XX PR 03-NOV-1999; 99US-0163280.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.

DR P-PSDB; AAG74109.

XX Claim 1; Page 2645-2646; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 1597 BP; 477 A; 323 C; 378 G; 418 T; 1 other;

Query Match 67.2%; Score 451.4; DB 22; Length 1597;
 Best Local Similarity 99.8%; Pred. No. 1.7e-69;
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGTAAAGAGATGGGTGAGATGATGAGGAACTCATATGATATATATTTGAA 215
 Db 1088 TGAGATTAAGGAGATGGGTGAGATGATGAGGAACTCATATGATATATTTGAA 1147
 QY 216 GATTATAGAGAGGAAATGCAAAATGACCAAAATTAACAATATGTCGGTGGAGC 275
 Db 1148 GATTATAGAGAGGAAATGCAAAATGACCAAAATTAACAATATGTCGGTGGAGC 1207
 QY 276 AAGACATCTTTGAAGGTGATGATTTGTAGTTAATCATATATTTGTATAGTAA 335
 Db 1208 AAGACATCTTTGAAGGTGATGATTTGTAGTTAATCATATATTTGTATAGTAA 1267
 QY 336 CCTGACTCAAAATTAATGACGCTGAAACGCTTACCAATCTTGAATTTGACCA 395
 Db 1268 CCTGACTCAAAATTAATGACGCTTGAACGCTTACCAATCTTGAATTTGACCA 1327
 QY 396 AGTGTCTTATATATGACATCTAATGTAAATCCAGAACTTGACCTCATCTTAAAT 455
 Db 1328 AGTGTCTTATATATGACATCTAATGTAAATCCAGAACTTGACCTCATCTTAAAT 1387
 QY 456 ATTATATGTAACTTCAATATGTGATTAATATGCTTCACAGTAATCTGAAAAA 515
 Db 1388 ATTTATGTGTAACTTCAATATGTGATTAATATGCTTCACAGTAATCTGAAAAA 1447
 QY 516 CTGATTTGTGATTTGAAGCTGCTTTCTATTACTTGAGTCTTGTACATACATCTTTT 575
 Db 1448 CTGATTTGTGATTTGAAGCTGCTTTCTATTACTTGAGTCTTGTACATACATCTTTT 1507
 QY 576 TATGAGCTATGAAATTAACATTTTAAACTGAA 608
 Db 1508 TATGAGCTATGAAATTAACATTTTAAACTGAA 1540

RESULT 7

ABO54273
 ID ABO54273 standard; cDNA; 1653 BP.

XX ABO54273;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HMLHW86 cDNA, SEQ ID NO:153.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 4q;
 KW gene; ss.

XX Homo sapiens.

XX MO200200677-A1.

XX PD 03-JAN-2002.

XX PR 07-JUN-2001; 2001MO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

QY 276 AAGACATCTTGAAGTCATGAGTTTGTAGTTTAACATCATATATTTGTATAGTGA 335
 DB 325 AAGACATCTTGAAGTCATGAGTTTGTAGTTTAACATCATATATTTGTATAGTGA 266
 QY 336 CCTGACTCAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCA 395
 DB 265 CCTGACTCAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCA 206
 QY 396 AGTGCTATATATATGACATGATTAATGTAATATCCAGAACTGGAGCTCCATGCTTAAAT 455
 DB 205 AGTGCTATATATATGACATGATTAATGTAATATCCAGAACTGGAGCTCCATGCTTAAAT 146
 QY 456 ATTATATGTAACATTCAAATGTGTGATTAATATGCTTCACAGTAATCTGAAAA 515
 DB 145 ATTATATGTAACATTCAAATGTGTGATTAATATGCTTCACAGTAATCTGAAAA 86
 QY 516 CTGATTTGTATGAAGCTGCTTTCTATTTACTTGAAGCTCTGTACATACATCTTTT 575
 DB 85 CTGATTTGTATGAAGCTGCTTTCTATTTACTTGAAGCTCTGTACATACATCTTTT 26
 QY 576 TATGAGCTATGAATAAATCAATTT 600
 DB 25 TATGAGCTATGAATAAATCAATTT 1
 RESULT 9
 ABR46008/c
 ID ABR46008 standard; cDNA: 481 BP.
 AC ABR46008;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE cDNA encoding colon tumour protein, SEQ ID NO.1559.
 XX
 KW Human; colon tumour; vaccine; colon cancer; immunogenic;
 KW Immunotherapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200212328-A2.
 PD 14-FEB-2002.
 XX
 PE 31-JUL-2001; 2001WO-US24218.
 XX
 PR 03-AUG-2000; 2000US-223283P.
 PR 28-MAR-2001; 2001US-279763P.
 PR 29-JUN-2001; 2001US-302051P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI King GE, Meagher MJ, Xu J, Secrist H;
 XX
 DR WPI; 2002-241739/29.
 XX
 PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
 PT for diagnosing, preventing, and treating colon cancer, and as markers
 PT for the progression of cancer -
 XX
 PS Claim 1; SEQ ID NO 1559; 147pp; English.
 XX
 CC The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABR44450-ABK65237 represent coding sequences of human colon tumour
 CC proteins of the invention.

CC Note: With the exception of SEQ ID NO 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.
 XX
 SO Sequence 481 BP; 149 A; 93 C; 64 G; 174 T; 1 other:
 Query Match 65.8%; Score 442.4; DB 24; Length 481;
 Best Local Similarity 99.6%; Pred. No. 5.7e-68;
 Matches 443; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 156 TCAGATAAGAGATGGGTGAGATGCATAGGAACTCATATGATATATATTTGA 215
 DB 445 TGAGATTAAGAGAGATGGGTGAGATGCATAGGAACTCATATGATATATTTGA 386
 QY 216 GATTATAGAAGAGAGAAATGCCAATGAGACACAAATTAAGTGTGGTGGAGC 275
 DB 385 GATTATAGAAGAGAGAAATGCCAATGAGACACAAATTAAGTGTGTGGTGGAGC 326
 QY 276 AAGACATCTTGAAGTCATGAGTTTGTAGTTTAACATCATATATTTGTATAGTGA 335
 DB 325 AAGACATCTTGAAGTCATGAGTTTGTAGTTTAACATCATATATTTGTATAGTGA 266
 QY 336 CCTGACTCAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCA 395
 DB 265 CCTGACTCAAAATATTAAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCA 206
 QY 396 AGTGCTATATATATGACATGATTAATGTAATATCCAGAACTGGAGCTCCATGCTTAAAT 455
 DB 205 AGTGCTATATATATGACATGATTAATGTAATATCCAGAACTGGAGCTCCATGCTTAAAT 146
 QY 456 ATTATATGTAACATTCAAATGTGTGATTAATATGCTTCACAGTAATCTGAAAA 515
 DB 145 ATTATATGTAACATTCAAATGTGTGATTAATATGCTTCACAGTAATCTGAAAA 86
 QY 516 CTGATTTGTATGAAGCTGCTTTCTATTTACTTGAAGCTCTGTACATACATCTTTT 575
 DB 85 CTGATTTGTATGAAGCTGCTTTCTATTTACTTGAAGCTCTGTACATACATCTTTT 26
 QY 576 TATGAGCTATGAATAAATCAATTT 600
 DB 25 TATGAGCTATGAATAAATCAATTT 1
 RESULT 10
 AAH83718/c
 ID AAH83718 standard; cDNA: 545 BP.
 AC AAH83718;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:1342.
 DE Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
 KW immunogenic; vaccine; ss.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200151513-A2.
 PD 19-JUL-2001.
 XX
 PE 16-JAN-2001; 2001WO-US01575.
 XX
 PR 14-JAN-2000; 2000US-0176722.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA;
 XX
 DR WPI; 2001-425866/45.
 XX
 PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to

PT treat and diagnose cancers, particularly ovarian cancer -
 XX Claim 5; Page 308; 338bp; English.
 XX
 CC AAH82377 to AAH83878 represent human ovarian tumour-associated
 CC polynucleotide sequences which encode ovarian tumour proteins. The
 CC ovarian tumour protein and polynucleotide sequences have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. The
 CC ovarian tumour proteins and polynucleotides can be used to inhibit
 CC the development of cancer, particularly ovarian cancer. They can also
 CC be used to diagnose the onset and progression of cancer.
 XX
 SQ Sequence 545 BP; 163 A; 112 C; 89 G; 177 T; 4 other;
 Query Match 65.8%; Score 442.2; DB 22; Length 545;
 Best Local Similarity 96.0%; Pred. No. 6.2e-68;
 Matches 453; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 129 GAATTTTCTGCTGCTTTCTCTGTTCAAGATTAAGAGATGGGTGATGATGAGGA 188
 DB 472 GGAATTCGCCNTTAGCGTGGCGCCGAGGTAAGAGATGGGTGATGATGAGGA 413
 QY 189 ACTGAATGCATTAATATATATTTGAAGATTATGAAGAGGAAATAGCAATGGACAC 248
 DB 412 ACTGAATGCATTAATATATATTTGAAGATTATGAAGAGGAAATAGCAATGGACAC 353
 QY 249 AATTAACAAATGCTGCTGCGTGGACGACAGACATCTTGAAGTCATGATTTGTTAGTT 308
 DB 352 AATTAACAAATGCTGCTGCGTGGACGACAGACATCTTGAAGTCATGATTTGTTAGTT 293
 QY 309 TAACATATATATTTGTAATAGTGAACCTGTACTCAAAATATAGCAGCTTGAACCTGG 368
 DB 292 TAACATATATATTTGTAATAGTGAACCTGTACTCAAAATATAGCAGCTTGAACCTGG 233
 QY 369 CTTACCAATCTGAAATTTGACACAAAGTCTTATATATGAGATCTATGTAATGAAATGC 428
 DB 232 CTTACCAATCTGAAATTTGACACAAAGTCTTATATATGAGATCTATGTAATGAAATGC 173
 QY 429 CAGAACTTGACCTCCATCGTTAAATATTTATGTATGATTCATCAATGTGTGCATTAAA 488
 DB 172 CAGAACTTGACCTCCATCGTTAAATATTTATGTATGATTCATCAATGTGTGCATTAAA 113
 QY 489 TATGCTTCCACAGTAAATCTGAAAACTGATTTGATTTGAAGCTGCTTTCTATTTA 548
 DB 112 TATGCTTCCACAGTAAATCTGAAAACTGATTTGATTTGAAGCTGCTTTCTATTTA 53
 QY 549 CTGAGCTCTGATCATATACATCTTTTATGAGCTATGAATAAATTTT 600
 DB 52 CTGAGCTCTGATCATATACATCTTTTATGAGCTATGAATAAATTTT 1
 RESULT 11
 AAS25073/C
 ID AAS25073 standard; cDNA; 439 BP.
 XX
 AC AAS25073;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human ovarian PCR-subtracted cDNA library clone #1254.
 XX
 XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
 KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
 KM primer; probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157207-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03733.
 XX

PR 04-FEB-2000; 2000US-0180403.
 PR 28-MAR-2000; 2000US-0192745.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Mannion J;
 XX
 DR WPI; 2001-488879/53.
 XX
 XX New polynucleotides encoding ovarian tumour proteins, useful for
 PT treating ovarian cancer, and as probes, primers, and markers of cancer
 PT progression -
 XX
 PS Example 1; page 305; 378bp; English.
 XX
 CC The invention comprises compositions used for the therapy and diagnosis
 CC of ovarian cancer. The compositions comprise one or more ovarian tumour
 CC proteins, their associated polynucleotides, or immunogenic portions of
 CC the proteins. The ovarian tumour polynucleotides and polypeptides are
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein. They are also useful for inhibiting the development of cancer in
 CC a patient with an ovarian tumour DNA or protein by incubating isolated
 CC T-cells allowing them to proliferate, and administering to the patient.
 CC The sequences can be used as markers for cancer, for example, to monitor
 CC ovarian cancer progression. Probes and primers are useful in nucleic acid
 CC hybridisation, in detecting the presence of complementary sequences in a
 CC given sample, for preparing mutant species and for preparing other
 CC genetic constructions. Sequences AAS22820-AAS25231 and AAS25238-AAS25549
 CC represent human ovarian tumour protein cDNA clones.
 XX
 SQ Sequence 439 BP; 142 A; 80 C; 60 G; 157 T; 0 other;
 Query Match 65.3%; Score 439; DB 22; Length 439;
 Best Local Similarity 100.0%; Pred. No. 2.2e-67;
 Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 162 AAGGAGATGGGTGAGATGATAGGAACTCAATGCACTATVNAATTTGAAGATTAT 221
 DB 439 AAGGAGATGGGTGAGATGATAGGAACTCAATGCACTATVNAATTTGAAGATTAT 380
 QY 222 AGAAGAGGAAATAGCAATGACACAAATTTCAAAATGTGTGTGGGAGGAGAC 281
 DB 379 AGAAGAGGAAATAGCAATGACACAAATTTCAAAATGTGTGTGGGAGGAGAC 320
 QY 282 TCTTTGAAGTCATGATTTGTTAGTTTACATCATATATTTGTAATAGGAAACCTGTA 341
 DB 319 TCTTTGAAGTCATGATTTGTTAGTTTACATCATATATTTGTAATAGGAAACCTGTA 260
 QY 342 CTCAAAATATATAGCAGCTTGAACCTGTACCAATCTGAAATTTGACACAGTGTG 401
 DB 259 CTCAAAATATATAGCAGCTTGAACCTGTGTACCAATCTGAAATTTGACACAGTGTG 200
 QY 402 TTTATATATGAGATCTATATTAATAATCCAGACTGTGACCTCCATCGTTAAATATTAT 461
 DB 199 TTTATATATGAGATCTATATTAATAATCCAGACTGTGACCTCCATCGTTAAATATTAT 140
 QY 462 GTGTAACATTCAAATGTGTGATTAATAATATGCTCCACAGATAAATCTGAAAAAGTATT 521
 DB 139 GTGTAACATTCAAATGTGTGATTAATAATATGCTCCACAGATAAATCTGAAAAAGTATT 80
 QY 522 TGTGATTTGAAGCTGCTTCTATTACTTGAAGCTTGTGATCATATCTTTTATGAG 581
 DB 79 TGTGATTTGAAGCTGCTTCTATTACTTGAAGCTTGTGATCATATCTTTTATGAG 20
 QY 582 CTATGAATTAACATTTT 600
 DB 19 CTATGAATTAACATTTT 1
 RESULT 12
 AAT29042/C
 ID AAT29042 standard; cDNA; 425 BP.
 XX

XX	AAI29042;
AC	
DT	12-OCT-2001 (first entry)
XX	
DE	Colon tumour related determined cDNA sequence for clone CS1-160.
XX	
KW	Human; immunotherapy; diagnosis; colon cancer; colon tumour;
XX	
KW	immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX	
OS	Homo sapiens.
PN	
XX	
PD	WO200149716-A2.
XX	
PF	29-DEC-2000; 2000OW-US35596.
XX	
PR	30-DEC-1999; 99US-0476296.
XX	
PR	10-JAN-2000; 2000US-0480321.
XX	
PR	15-FEB-2000; 2000US-0504629.
XX	
PR	06-MAR-2000; 2000US-0519444.
XX	
PR	19-MAY-2000; 2000US-0575251.
XX	
PR	29-JUN-2000; 2000US-0609448.
XX	
PR	28-AUG-2000; 2000US-0649811.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Xu J., Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
XX	
PI	King GE, Wang T, Jiang Y;
XX	
DR	WPI: 2001-441847/47.
XX	
PT	Colo tumor associated proteins and nucleic acids useful for the
XX	
PT	prevention, diagnosis and treatment of colonic cancer -
XX	
PS	Claim 25; Page 291; 472pp; English.
XX	
XX	The present invention describes colon tumour associated proteins (I) and
CC	the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC	(I) and (II) can be used in gene therapy and vaccine production. (I) and
CC	(II) may be used in the prevention, diagnosis and treatment of diseases
CC	associated with inappropriate colon tumour associated protein (TCAP)
CC	expression, such as colonic cancer. For example, (I) and (II) may be
CC	used to treat disorders associated with decreased expression by
CC	rectifying mutations or deletions in a patient's genome that affect the
CC	activity of TCAPs by expressing inactive proteins or to supplement the
CC	patients own production of them. Additionally, (II) may be used to
CC	produce the TCAP proteins, by inserting the nucleic acids into a host
CC	cell culturing the cell to express the protein. (II) and its
CC	complementary sequences may also be used as DNA probes in diagnostic
CC	polymerase chain reaction (PCR) and hybridisation assays to detect and
CC	quantitate the presence of similar nucleic acids in samples, and
CC	therefore which patients may be in need of restorative therapy. (I) may
CC	also be used as antigens in the production of antibodies against TCAPs
CC	and in assays to identify modulators of TCAP expression and activity.
CC	Anti-(I) antibodies and antagonists may also be used to down regulate
CC	TCAP expression and activity. The anti-(I) antibodies may also be used
CC	as diagnostic agents for detecting the presence of TCAPs in samples
CC	(e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
CC	and AAI24494 to AAM24523 represent nucleotide and amino acid sequences
CC	given in the exemplification of the present invention.
XX	
SQ	Sequence 425 BP; 132 A; 83 C; 59 G; 151 T; 0 other:
XX	
Query Match	61.8%; Score 415.4; DB 22; Length 425;
Best Local Similarity	99.8%; Pred. No. 2.7e-63;
Matches 416; Conservative	0; Mismatches 1; Indels 0; Gaps 0
OY	156 TCAGATTAAGGAGATGGCGTGCATGTGGGAACCAATGCATTACTATATAATTGAA 215
db	417 TCAGTAAAGGAGATGGCGTGAGTGTGATGGGCACTCAATGCTATATATATTGAA 358
OY	216 GATTTAGAAGAAGGAATAGCAATGACACAATTTACAAATGTGTGTGCTGGGACG 275

Db	357	GATTATAGAGAGAGGGAATATGCAATATGACACAAATTACAAATGTGTGTGGTGGACG	298
Qy	276	AAGACATCTTTGAAGGTCATGAGTTTGTATGTTTAAACATCATATATTTGTAATAGTGA	335
Db	297	AAGACATCTTGAAGGTCATGAGTTTGTATGTTTAAACATCATATATTTGTAATAGTGA	238
Qy	336	CCTGTACTCAAAATATATAGCAGCTGAAACTGGCTTTACCACATCTTTGAATTTGACACA	395
Db	237	CCTGTACTCAAAATATATAGCAGCTGAAACTGGCTTTACCACATCTTTGAATTTGACACA	178
Qy	396	AGTGTCTTATATATGACATCTAATGAAATCCACAACTTGAGCTCATCTGTAAATTT	455
Db	177	AGTGTCTTATATATGACATCTAATGAAATCCACAACTTGAGCTCATCTGTAAATTT	118
Qy	456	ATTATATGTGTACATTCATAAATGTGTGCATTAATAATATGCTTCCACAGTAAATCTGA	515
Db	117	ATTATATGTGTACATTCATAAATGTGTGCATTAATAATATGCTTCCACAGTAAATCTGA	58
Qy	516	CTGATTTGTGTATGAAGCTGCCCTTCTATTTACTTCTGAGCTTGTACATACATCTT	572
Db	57	CTGATTTGTGTATGAAGCTGCCCTTCTATTTACTTCTGAGCTTGTACATACATCTT	1
RESULT 13			
ID	ABK36208	standard; cDNA: 1439 BP.	
AC	ABK36208;		
XX	08-MAY-2002 (first entry)		
DT	CDNA sequence #599 encoding novel human secreted protein.		
XX			
DE	Human secreted protein; hyperproliferative disorder; autoimmune disorder;		
XX	Immune deficiency disorder; blood disorder; inflammatory disorder;		
KM	Infectious disorder; allergic condition; neurodegenerative disorder;		
KW	Liver fibrosis; congenital disorder; gene therapy; antitubercular;		
XX	tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
XX	WO200177289-A2.		
PN	18-OCT-2001.		
XX			
PD	29-MAR-2001; 2001WO-US10232.		
XX			
PF	06-APR-2000; 2000US-195605P.		
XX			
PR	(GENY) GENETICS INST INC.		
PA			
XX	Jeabs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;		
PI	Mebergs D, Treacy M, Apostino MJ, Bowman MR, Spaulding V, Wong GS;		
PI	Clark HF, Reichel K, Howes SH, Resnick RJ, Gulikova K, Graham JR;		
DR	WPI: 2002-179322/23.		
XX			
XX	Claim 1; Page 384; 393pp: English.		
PS			
XX			
CC	The present invention relates to the isolation of novel cDNA sequences		
CC	which encode human secreted proteins. The cDNA sequences have been		
CC	derived from a variety of human tissues. The invention also provides		
CC	a method for producing proteins from these polynucleotide sequences.		
CC	The proteins are useful for identifying compounds that modulate their		
CC	activity and production. The sequences of the invention are		
CC	useful for treating diseases such as hyperproliferative disorders		
CC	(e.g. cancer), immune deficiency disorders (e.g. severe combined		

CC Immunodeficiency (SCID), autoimmune disorders (e.g. multiple
 CC scleroderma), blood disorders (e.g. thrombocytopenia), inflammatory
 CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),
 CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.
 CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.
 CC haemophilia), and tumours. The polynucleotide sequences of the
 CC invention are also useful in gene therapy. ABK35610-ABK36232 represent
 CC the cDNA sequences of the invention that encode for novel human
 CC secreted proteins.

SQ Sequence 1439 BP; 417 A; 301 C; 351 G; 370 T; 0 other;

Query Match 53.8%; Score 361.4; DB 24; Length 1439;
 Best Local Similarity 99.7%; Pred. No. 5,9e-54;
 Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGATAAAGAGATGGTGGATGATGATAGGAGCAATCAGTACTATATATTTGAA 215
 DB 1077 TGAAGATAAAGAGATGGTGGATGATGATAGGAGCAATCAGTACTATATATTTGAA 1136
 QY 216 GATTATAGAAGAGAGAAATAGCAATGACACAAATGCTGTGCGTGGAGC 275
 DB 1137 GATTATAGAAGAGAGAAATAGCAATGACACAAATGCTGTGCGTGGAGC 1196
 QY 276 AAGACATCTTTGAGAGCTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGTGA 335
 DB 1197 AAGACATCTTTGAGAGCTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGTGA 1256
 QY 336 CCTGTACTCAAAATATAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 395
 DB 1257 CCTGTACTCAAAATATAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 1316
 QY 396 AGTGTCTTATATATGAGATCTATATGTAATAATCCAGAACTGGACTCCATCGTTAAAT 455
 DB 1317 AGTGTCTTATATATGAGATCTATATGTAATAATCCAGAACTGGACTCCATCGTTAAAT 1376
 QY 456 ATTATATGTATACATTCAAATGTGTGTCATTAAATATGCTTCCACAGTAAATCTGAAAA 515
 DB 1377 ATTATATGTATACATTCAAATGTGTGTCATTAAATATGCTTCCACAGTAAATCTGAAAA 1436
 QY 516 CTG 518
 DB 1437 CTG 1439

RESULT 14

AAC69778
 ID AAC69778 standard; cDNA; 536 BP.

XX AAC69778;

DT 30-JAN-2001 (first entry)

XX Human breast tumour antigen partial coding sequence #77.

XX Cytostatic; vaccine; human; breast tumour; antigen; breast cancer; ss.

OS Homo sapiens.

XX WO200052165-A2.

PN 08-SEP-2000.

XX 29-FEB-2000; 2000WO-US05431.

PR 04-MAR-1999; 99US-0262505.

PR 19-MAR-1999; 99US-0272886.

XX 17-SEP-1999; 99US-0396313.

PA (CORI-) CORIXA CORP.

XX Lodes MJ;

DR WP1: 2000-572184/53.
 XX Breast tumor antigen polypeptides and polynucleotides, useful for
 PT manufacturing vaccines and compositions for treating, diagnosing, and
 PT monitoring breast cancer.

PS Claim 16; Page 128; 140pp; English.

CC The present invention relates to immunogenic portions of new human
 CC breast tumor antigens (AAB28183-B28214) and their coding sequences
 CC (AAC69645-C69804). The breast tumor antigen polypeptides of the present
 CC invention and their coding sequences are useful for inhibiting the
 CC development of breast cancer in a patient. The breast tumor antigen
 CC polypeptides and polynucleotides may be used in vaccines and
 CC pharmaceutical compositions for treating breast cancer, and for
 CC diagnosing and monitoring the cancer. The present sequence is a coding
 CC sequence for the immunogenic portion for one such human breast cancer
 CC tumor antigen.

SQ Sequence 536 BP; 172 A; 78 C; 127 G; 159 T; 0 other;

Query Match 51.4%; Score 345.4; DB 21; Length 536;
 Best Local Similarity 99.7%; Pred. No. 3.3e-51;
 Matches 346; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGATAAAGAGATGGTGGATGATGATAGGAGCAATCAGTACTATATATTTGAA 215
 DB 113 TGAAGATAAAGAGATGGTGGATGATGATAGGAGCAATCAGTACTATATATTTGAA 172
 QY 216 GATTATAGAAGAGAGAAATAGCAATGACACAAATGCTGTGCGTGGAGC 275
 DB 173 GATTATAGAAGAGAGAAATAGCAATGACACAAATGCTGTGCGTGGAGC 232
 QY 276 AAGACATCTTTGAGAGCTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGTGA 335
 DB 233 AAGACATCTTTGAGAGCTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGTGA 292
 QY 336 CCTGTACTCAAAATATAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 395
 DB 293 CCTGTACTCAAAATATAGCAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA 352
 QY 396 AGTGTCTTATATATGAGATCTATATGTAATAATCCAGAACTGGACTCCATCGTTAAAT 455
 DB 353 AGTGTCTTATATATGAGATCTATATGTAATAATCCAGAACTGGACTCCATCGTTAAAT 412
 QY 456 ATTATATGTATACATTCAAATGTGTGTCATTAAATATGCTTCCACAGT 502
 DB 413 ATTATATGTATACATTCAAATGTGTGTCATTAAATATGCTTCCACAGT 459

RESULT 15

AAS25360
 ID AAS25360 standard; cDNA; 507 BP.

XX AAS25360;

DT 07-NOV-2001 (first entry)

XX Human ovarian PCR-subtracted cDNA library clone #1445.

XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;

XX gene therapy; cytosolic; T-cell expansion; nucleic acid hybridisation;
 KW primer; probe.

OS Homo sapiens.

XX WO200157207-A2.

PN 09-AUG-2001.

PR 05-FEB-2001; 2001WO-US03733.

XX 04-FEB-2000; 2000US-0180403.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 19:45:25 ; Search time 629.416 Seconds
(without alignments)
17291.221 Million cell updates/sec

Title: US-09-919-497-32

Perfect score: 672

Sequence: 1 aatactattcttctttt.....tttttttttttttgaga 672

Scoring table: IDENTITY, NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

EST: *

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hnc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hnc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pla:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	503.8	75.0	554	9	A1027882
C 2	452.8	67.4	471	10	AM769100
C 3	451.4	67.2	501	12	BG292290
C 4	451.4	67.2	573	13	BI438810
C 5	451.4	67.2	592	13	BM272151
C 6	451.4	67.2	604	10	AM473340

C 7	451.4	67.2	614	9	A1624711
C 8	451.4	67.2	627	9	A1955311
C 9	451.4	67.2	675	9	A1978632
C 10	451.4	67.2	709	14	BM981479
C 11	451.4	67.0	768	9	A1571882
C 12	449.8	66.9	513	9	AA502398
C 13	449.8	66.9	471	9	A1038566
C 14	449.8	66.9	481	13	BM510869
C 15	449.8	66.9	535	10	BE672486
C 16	449.8	66.9	601	10	AM513282
C 17	449.8	66.9	683	9	A1962963
C 18	449.8	66.9	690	10	AM188163
C 19	449.4	66.9	511	9	AA424571
C 20	449.4	66.9	561	9	AU155306
C 21	449.4	66.9	570	10	AM780425
C 22	449.4	66.9	684	10	AM004850
C 23	449.4	66.9	686	9	A1628538
C 24	449.4	66.9	764	10	AM103479
C 25	449.4	66.9	830	9	A1718512
C 26	449.2	66.8	506	9	A1285353
C 27	448.8	66.8	520	10	AM473388
C 28	448.8	66.8	618	10	AM272287
C 29	448.8	66.8	694	10	AM129970
C 30	448.4	66.7	542	9	A1241214
C 31	448.4	66.7	598	14	BO416873
C 32	448.4	66.7	603	14	BO778413
C 33	448.2	66.7	612	14	BO778448
C 34	448.2	66.7	551	9	A1024785
C 35	447.8	66.6	707	10	AM007667
C 36	447.4	66.6	516	9	A1285185
C 37	447.2	66.5	681	9	A1697298
C 38	447.2	66.5	752	9	A1745085
C 39	443.4	66.0	692	9	A1721217
C 40	441.8	65.7	468	9	AA443854
C 41	440.2	65.5	605	10	AM873982
C 42	439.8	65.4	453	9	A1074078
C 43	437.8	65.1	579	9	AA130871
C 44	437.4	65.1	486	9	A1824537
C 45	437.2	65.1	518	10	AM075441

ALIGNMENTS

RESULT 1
A1027882/c
LOCUS
DEFINITION
ow52e04.x1 Soares-parathyroid.tumor.NbHPA Homo sapiens cDNA clone IMAGE:1650462.3' similar to gb:U14758 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED PROTEIN GA733-2 (HUMAN);, mRNA sequence.

ACCESSION
A1027882
VERSION
A1027882.1 GI:3246581

KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 554)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo,
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html
Insert Length: 1292 Std Error: 0.00
Seq primer: -40ml3 fwd. EY from Amersham

```

FEATURES
    source
        High quality sequence stop: 455.
        Location/Qualifiers
            1..554
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="IMAGE:1650462"
                /clone_11b="Soares.parathyroid_tumor_NbHPA"
                /tissue_type="parathyroid tumor"
                /dev_stage="adult"
                /lab_host="DH10B (ampicillin resistant)"
                /note="Organ: parathyroid gland; Vector: p773D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - 0190(drr) primer"
                [5'-GTGTCACATCTGATGGGAGCGCCGACCAATTTTTTTTTTTTTTTTTTTT
                TTTT-3'].. double-stranded cDNA was size selected, ligated
                to Eco RI adaptters (Pharmacia), digested with Not I and
                cloned into the Not I and Eco RI sites of a modified p7733
                vector (Pharmacia). Library went through one round of
                normalization to a Cot = 5. Library constructed by Bento
                Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                adenomas was kindly provided by Dr. Stephen Marx, National
                Institute of Diabetes and Digestive and Kidney Diseases,
                NIH."

```

Query Match	75.0%	Score 503.8	DB 9	Length 554
Best Local Similarity	99.4%	Pred. No. 9e-64		
Matches 516	Conservative 0	Mismatches 2	Indels 1	Gaps 1
QY	. 1	AATATACCTTTTCTTTTATTTAAATAGCTCTTAAATGTGGGAAAAATTATTCCT	60	
Db	523	AATATACCTTTTCTTTTATTTAAATAGCTCTTAAATGTGGGAAAAATTATTCCT	464	
QY	61	GTGTTCCCTTAATTCATTTTATTTAACTACTATTTTACGAATGACCAAAAGATTGAANA	120	
Db	463	GTGTTCCCTTAATTCATTTTATTTAACTACTATTTTACGAATGACCAAAAGATTGAANA	404	
QY	121	ATTATTTGAATTTTCTGTCGCTTTTCTGTTCCGTTTCAGATTAAGAGATGGGTGATG	180	
Db	403	ATTATTTGAATTTTCTGTCGCTTTTCTGTTCCGTTTCAGATTAAGAGATGGGTGATG	345	
QY	181	CATAGGGAAGCTCAATGCATTAACCTATATTAATTGAAGATTATAGAGAAAGCAATATAGCA	240	
Db	344	CATAGGGAAGCTCAATGCATTAACCTATATTAATTGAAGATTATAGAGAAAGCAATATAGCA	285	
QY	241	ATGGACACAATTTACAAATGTGTGTGTGGCGTGGGACCAAGACATCTTTGAAGGTGATGAGTT	300	
Db	284	ATGGACACAATTTACAAATGTGTGTGTGGCGTGGGACCAAGACATCTTTGAAGGTGATGAGTT	225	
QY	301	TGTTAGTTTAAACATATATTTTGTATATAGTGAACCTGTACTCAAAATATATAGCAGCTT	360	
Db	224	TGTTAGTTTAAACATATATTTTGTATATAGTGAACCTGTACTCAAAATATATAGCAGCTT	165	
QY	361	GAAACTGGCTTTACCAATCTTGAAAATTTGACCACAAGTGTCTTATATATATGAGATCTAAT	420	
Db	164	GAAACTGGCTTTACCAATCTTGAAAATTTGACCACAAGTGTCTTATATATATGAGATCTAAT	105	
QY	421	GTAATAATCCAGAACCTTGGACATCCATCGTTAAATTTATTTATGTGTAAACATTCATATGTCT	480	
Db	104	GTAATAATCCAGAACCTTGGACATCCATCGTTAAATTTATTTATGTGTAAACATTCATATGTCT	45	
QY	481	GCATTAATATATGCTTCGACAGATTAATATCGAAAAACTGA 519		
Db	44	GCATTAATATATGCTTCGACACCAAAATCTGAAAAAACTGA 6		

RESULT 2

AM769100/c

LOCUS

DEFINITION

h154802.x1 NCI_CGAP_C014 Homo Sapiens CDNA clone IMAGE:3004970 3'

Similar to gb:U14738 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED

FEATURES	SOURCE
JOURNAL COMMENT	PROTEIN GA733-2 (HUMAN); mRNA sequence.
REFERENCE TITLES	AM769100 AM769100.1 GI:7701127 EST.
AUTHORS	human. Homo sapiens
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
URL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
DESCRIPTION	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN, send email to: InfoImage.Lnl.nsl.gov Seq primer: -40UP from Gabc High quality sequence stop: 395.
LOCATION/QUALIFIERS	I..471

		libseq_type="moderately-differentiated adenocarcinoma"	alias host="DHD08"	/note="Organ: colon; Vector: PCMV-SpRt6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Life Technologies catalog #:
BASE COUNT	148 a	87 c	62 g	174 t
ORIGIN				11531-019"
Query Match		67.4%	Score 452.8;	DB 10; Length 471;
Best Local Similarity		99.6%	Pred. No.2.2e-56;	
Matches 454;	Conservatively	0;	Mismatches 2;	Indels 0; Gaps 0;
Dy	153	GTTCAGATTAAGAGGATGGGTGAGATGCATAGGGAACCTCATGCTATATATATTT	212	
Db	459	GCTTGAGATTAAGGAGATGGGTGAGATGCATAGGGAACCTCATGCTATATATATTT	400	
Qy	213	GAACTTTTAAAGAAAGGAAATAGCAATAGACCAATATACAAATCTGTGTGGTGGG	272	
Db	399	GAACTTTTAAAGAAAGGAAATAGCAATAGACCAATATACAAATCTGTGTGGTGGG	340	
Qy	273	ACGAAGACATCTTTGAAAGGTCATGAGTTTGTAGTTTAACTCATATATTTTGTATAGTG	332	
Db	339	ACGAAGACATCTTTGAAAGGTCATGAGTTTGTAGTTTAACTCATATATTTTGTATAGTG	280	
Qy	333	AAACCTGTACCTCAAAATATTAAGCAGCTTGAACCTGGCTTACCACTCTTGAATTTTGACC	392	
Db	279	AAACCTGTACCTCAAAATATTAAGCAGCTTGAACCTGGCTTACCACTCTTGAATTTTGACC	220	
Qy	393	ACAAGTGTCTATATATATGTCAGATCTTAATGTAAATCCAGAACTTGGACTCCATCGTTAAA	452	
Db	219	ACAAGTGTCTATATATATGTCAGATCTTAATGTAAATCCAGAACTTGGACTCCATCGTTAAA	160	
Qy	453	ATTATTTTGTGTACATCTTCAATGTGTGCATTAATATGCTTCCACAGTAAATCTGAA	512	
Db	159	ATTATTTTGTGTACATCTTCAATGTGTGCATTAATATGCTTCCACAGTAAATCTGAA	100	
Qy	513	AAACGATTTTGTGATGTAAGACCTGGCTTCTATTTACTTGAAGCTCTGTACATACACTACTT	572	
Dy	99	AAACGATTTTGTGATGTAAGACCTGGCTTCTATTTACTTGAAGCTCTGTACATACACTACTT	40	
Qy	573	TTTTATGAGCTATGAAATAAACATTTTAAACTGAA	608	

QY 216 GATTATAGAGAGGAAATAGCAATATGACACAAATATGCTGCGTGGAGC 275
 |||||||
 Db 400 GATTATAGAGAGGAAATAGCAATATGACACAAATATGCTGCGTGGAGC 341
 |||||||
 QY 276 AAGACATCTTTGAAGTCATGAGTTGTTAGTTTACATCATATATTGTAATAGTGA 335
 |||||||
 Db 340 AAGACATCTTTGAAGTCATGAGTTGTTAGTTTACATCATATATTGTAATAGTGA 281
 |||||||
 QY 336 CCTGTACTCAAAATATAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGC 395
 |||||||
 Db 280 CCTGTACTCAAAATATAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGC 221
 |||||||
 QY 396 AGTGTCTTATATATGACATCTATATGTAATATGCAAGCTTGAAGCTTGAAGCTTGAAGT 455
 |||||||
 Db 220 AGTGTCTTATATATGACATCTATATGTAATATGCAAGCTTGAAGCTTGAAGCTTGAAGT 161
 |||||||
 QY 456 ATTATGTGTACATCTCAAAATGTCGATTAATATGCTTCCACAGTAATATCTGAAAA 515
 |||||||
 Db 160 ATTATGTGTACATCTCAAAATGTCGATTAATATGCTTCCACAGTAATATCTGAAAA 101
 |||||||
 QY 516 CTGATTTGATGTAAGAGCGCTTCTATTTACTGAGCTTGTACATCATATCTTTT 575
 |||||||
 Db 100 CTGATTTGATGTAAGAGCGCTTCTATTTACTGAGCTTGTACATCATATCTTTT 41
 |||||||
 QY 576 TATGAGCTATGAATTAACATTTTAACTGAA 608
 |||||||
 Db 40 TATGAGCTATGAATTAACATTTTAACTGAA 8
 |||||||

RESULT 5
 BM272151/c
 LOCUS 592 bp mRNA linear EST 20-DEC-2001
 DEFINITION 1g39g02.x1 HR85 Islet Homo sapiens cDNA 3' similar to SM:G332.HUMAN
 P16422 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED PROTEIN GAV33-2
 PRECURSOR: mRNA sequence.

ACCESSION BM272151
 VERSION BM272151
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Melton,D., Brown,D., Kenty,G., Pernutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Sceane,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Riltter,E., Ronko,I., Bennett,D., Cardenas
 , M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
 Jackson,Y., and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)

TITLE JOURNAL
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohpc.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@um.wustl.edu)
 Seq primer: -400P from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers
 1. 592

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HR85 Islet"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:"

NOTI: Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact Information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Pernutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@umcgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692.

BASE COUNT 188 a 124 c 77 g 203 t
 ORIGIN

Query Match 67.2%; Score 451.4; DB 13; Length 592;
 Best Local Similarity 99.8%; Pred. No. 3e-56;
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 156 TCAGATTAAGAGAGTGGTGCATGATGATGAGGAACTCAATGATATATATTTGAA 215
 |||||||
 Db 455 TGAGATTAAGAGAGTGGTGCATGATGATGAGGAACTCAATGATATATATTTGAA 396
 |||||||
 QY 216 GATTATAGAGAGGAAATAGCAATATGACACAAATATGTCGCGTGGAGC 275
 |||||||
 Db 395 GATTATAGAGAGGAAATAGCAATATGACACAAATATGTCGCGTGGAGC 336
 |||||||
 QY 276 AAGACATCTTTGAAGTCATGAGTTGTTAGTTTACATCATATATTGTAATAGTGA 335
 |||||||
 Db 335 AAGACATCTTTGAAGTCATGAGTTGTTAGTTTACATCATATATTGTAATAGTGA 276
 |||||||
 QY 336 CTGTACTCAAAATATAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGC 395
 |||||||
 Db 275 CTGTACTCAAAATATAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGC 216
 |||||||
 QY 396 AGTGTCTTATATATGACATCTATATGTAATATGCAAGCTTGAAGCTTGAAGCTTGAAGT 455
 |||||||
 Db 215 AGTGTCTTATATATGACATCTATATGTAATATGCAAGCTTGAAGCTTGAAGCTTGAAGT 156
 |||||||
 QY 456 ATTATGTGTACATCTCAAAATGTCGATTAATATGCTTCCACAGTAATATCTGAAAA 515
 |||||||
 Db 155 ATTATGTGTACATCTCAAAATGTCGATTAATATGCTTCCACAGTAATATCTGAAAA 96
 |||||||
 QY 516 CTGATTTGATGTAAGAGCGCTTCTATTTACTGAGCTTGTACATCATATCTTTT 575
 |||||||
 Db 95 CTGATTTGATGTAAGAGCGCTTCTATTTACTGAGCTTGTACATCATATCTTTT 36
 |||||||
 QY 576 TATGAGCTATGAATTAACATTTTAACTGAA 608
 |||||||
 Db 35 TATGAGCTATGAATTAACATTTTAACTGAA 3

RESULT 6
 AM473340/c
 LOCUS 604 bp mRNA linear EST 24-FEB-2000
 DEFINITION hb03h06.x1 NCI-GAP_Gas4 Homo sapiens cDNA clone IMAGE:2882171 3'
 similar to gb:X14758 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED
 PROTEIN GAV33-2 (HUMAN); mRNA sequence.

ACCESSION AM473340
 VERSION AM473340
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS 1 (bases 1 to 604)

TITLE JOURNAL
 COMMENT NCI-GAP http://www.ncbi.nlm.nih.gov/ncigap.
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

LOCUS	AI955311/c	627 bp	mRNA	linear	EST 09-MAR-2000
DEFINITION	AI955311				
DESCRIPTION	wx22ell.x1 NCI-CGAP pit1 Homo sapiens cDNA clone IMAGE:2545388 3'				
REFERENCE	1 (bases 1 to 627)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLND at: www.bio.lnl.gov/bbrp/image/image.html Insert length: 1680 Std Error: 0.00 Seq primer: -400p from Gibco High quality sequence stop: 410. Location/Qualifiers				
FEATURES	source	1..627			
	organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:2545388"				
	/clone_id="NCI CGAP Pit1"				
	/tissue_type="four pooled pituitary adenomas"				
	/lab_host="DH10B"				
	/note="Organ: brain; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies."				
BASE COUNT	201 a 131 c 79 g 213 t				
ORIGIN					
Query Match	67.2%	Score 451.4:	DB 9:	Length 627:	
Best Local Similarity	99.8%:	Pred. No. 2.9e-56:			
Matches 452; Conservative	0; Mismatches 1;	Indels 0;	Gaps 0;		
156	TCAGATAAAGGAGATGGGAGATGATGATGAGGAAACCATGATACATATATATATTTGAA	215			
456	TGAGATAAAGGAGATGGGAGATGATGATGAGGAAACCATGATACATATATATATTTGAA	397			
216	GATTTATAGAAGGAGGAAATAGCAATGAGACACAAATTCAAATGTGTGCGTGGGACG	275			
396	GATTTATAGAAGGAGGAAATAGCAATGAGACACAAATTCAAATGTGTGCGTGGGACG	337			
276	AAGACATCTTTGAAGTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGGAA	335			
336	AAGACATCTTTGAAGTCATGAGTTGTTAGTTTAACATCATATATTTGTAATAGGAA	277			
336	CCTGTACTCAAAATATATACGAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA	395			
276	CCTGTACTCAAAATATATACGAGCTTGAACCTGGCTTACCAATCTTGAATTTGACCACA	217			
396	AGTGTCTTATATATGCAAGATCTTAATGTAATATCCAACTGGACATCCATCGTTAAAT	455			
216	AGTGTCTTATATATGCAAGATCTTAATGTAATATCCAACTGGACATCCATCGTTAAAT	157			
456	ATTATATGTTAATCAATGATGTGATTTAATATGCTTCCACAGTAAATCTGAAAAA	515			
156	ATTATATGTTAATCAATGATGTGATTTAATATGCTTCCACAGTAAATCTGAAAAA	97			
516	CTGATTTTGATTAAGAAAGCTGCTTCTATTTACTGTTGAGTCTTGATCATACATACCTTTT	575			

Db	96	CTGATTGTGATGTAAGACCTGCTTTCATTTACTTACTGAGCTTGACATACATCTTTT	37
QY	576	TATGAGCTATGAATTAACATTTTAACGTAA 608	
Db	36	TATGAGCTATGAATTAACATTTTAACGTAA 4	
RESULT 9	AI978632/c		
LOCUS	AI978632	675 bp	mRNA
DEFINITION	wi7fe02.x1 NCI-CGAP-Ucl Homo sapiens cDNA clone IMAGE:2491802 3'		
ACCESSION	AI978632		
VERSION	AI978632.1	GI:5803662	
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 675)		
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@b-remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/db/rrp/image/image.html Insert length: 1133 Std Error: 0.00 Seq primer: -40up from G100 High quality sequence stop: 336.		
FEATURES	Location/Qualifiers		
source	1..675		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2491802"		
	/clone_11b="NCI CGAP ut1"		
	/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"		
	/lab_host="DH10B"		
	/note="Organ: uterus; Vector: pCMV-Sport6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"		
BASE COUNT	210 a 141 c 89 g 234 t	1 others	
ORIGIN			
Query Match	67.2%;	Score 451.4;	DB 9; Length 675;
Best Local Similarity	99.8%;	Pred. No. 2.8e-56;	
Matches 452;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	156	TCAGTAAAGGAGATGCGTGAGATGATGATGGGAAGCTCATGCATATATATATATTGAA 215	
Db	460	TCAGTAAAGGAGATGCGTGAGATGATGATGGGAAGCTCATGCATATATATATTGAA 401	
QY	216	GATTATAGAGAAGGAATAGCAATGAGACACAAATTCACAAATGCTGCGTGGAGC 275	
Db	400	GATTATAGAGAAGGAATAGCAATGAGACACAAATTCACAAATGCTGCGTGGAGC 341	
QY	276	AAGACATCTTTGAAGCTCATGAGTTGTTAGTTAATCATATATTTGTATTAAGTAA 335	
Db	340	AAGACATCTTTGAAGCTCATGAGTTGTTAGTTAATCATATATTTGTATTAAGTAA 281	
QY	336	CGTATCTCAAAATATTAACAGCTGAACCTGACCTTACCATCTTGAATTTGACACA 395	

Db 280 CCTGACTCAAAATATTAAGACGCTTGAAACGTGGCTTTACCAATCTTGAAATTTGACCA 221
 Qy 396 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGGAGTCATGCTTAAAT 455
 Db 220 AGTGCTATATATGACAGATCTAATGTAAATCCAGAACTGGAGTCATGCTTAAAT 161
 Qy 456 ATTTATGTGTACATTCAAATGTGTGCAATTAATATGCTTCACAGTAATTCGAAAA 515
 Db 160 ATTTATGTGTACATTCAAATGTGTGCAATTAATATGCTTCACAGTAATTCGAAAA 101
 Qy 516 CTGATTTGTGATGAACCTGCTTCTATTTACTGTGATGCTGTGACATACATCTTTT 575
 Db 100 CTGATTTGTGATGAACCTGCTTCTATTTACTGTGATGCTGTGACATACATCTTTT 41
 Qy 576 TATGAGCTATGAAATTAACATTTTAACTGAA 608
 Db 40 TATGAGCTATGAAATTAACATTTTAACTGAA 8
 RESULT 10
 BM981479/c 709 bp mRNA linear EST 21-MAR-2002
 LOCUS UI-CF-EN1-adh-h-02-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
 DEFINITION UI-CF-EN1-adh-h-02-0-UI 3', mRNA sequence.
 ACCESSION BM981479
 VERSION BM981479.1 GI:19604010
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 709)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs., Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 505-573, >(CAA)n#Simple_repeat
 Seq primer: M13 FORWARD
 POLY-A=yes. Location/Qualifiers
 1. 709
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: Lung; Vector: pRT3-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-CF-EN1 is a normalized cDNA library containing the
 following tissue(s): Primary Lung Cystic Fibrosis
 Epithelial Cells. The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not

I, and cloned directionally into pRT3-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (df)18 tail. The
 sequence tag for this library is CTGCTCAGGT.
 TAG_Lib-UI-CF-EN1
 TAG_TISSUE-Human Lung Epithelial Cell Lines untreated LPS
 6hr to LPS 24h
 TAG_SEQ-CTGCTCAGGT"
 BASE COUNT 221 a 146 c 91 g 251 t
 ORIGIN
 Query Match 67.2%; Score 451.4; DB 14; Length 709;
 Best Local Similarity 99.8%; Pred. No. 2.7e-56;
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 156 TCAGATAAAGGATGGGATGATGCATATGGAAGCAATGATATATTAATTTGAA 215
 Db 469 TGAGATTAAGGATGGGATGATGCATATGGAAGCAATGATATATTAATTTGAA 410
 Qy 216 GATTATAGAGAGGAGGAATAGCAATGAGACACAATTTACAATGTGTGCGTGGAGC 275
 Db 409 GATTATAGAGAGGAGGAATAGCAATGAGACACAATTTACAATGTGTGCGTGGAGC 350
 Qy 276 AAGACATCTTTGAGGATGATGATGTTAGTTTAACATCATATATTTGTAATAGTGANA 335
 Db 349 AAGACATCTTTGAGGATGATGATGTTAGTTTAACATCATATATTTGTAATAGTGANA 290
 Qy 336 CCTGACTCAAAATATTAAGCAGCTTGAAGCTGCTTTACCAATCTTGAATTTGACCA 355
 Db 289 CCTGACTCAAAATATTAAGCAGCTTGAAGCTGCTTTACCAATCTTGAATTTGACCA 230
 Qy 396 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGGAGTCATGCTTAAAT 455
 Db 220 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGGAGTCATGCTTAAAT 170
 Qy 456 ATTTATGTGTACATTCAAATGTGTGCAATTAATATGCTTCACAGTAATTCGAAAA 515
 Db 160 ATTTATGTGTACATTCAAATGTGTGCAATTAATATGCTTCACAGTAATTCGAAAA 110
 Qy 516 CTGATTTGTGATGAACCTGCTTCTATTTACTGTGATGCTGTGACATACATCTTTT 575
 Db 109 CTGATTTGTGATGAACCTGCTTCTATTTACTGTGATGCTGTGACATACATCTTTT 50
 Qy 576 TATGAGCTATGAAATTAACATTTTAACTGAA 608
 Db 49 TATGAGCTATGAAATTAACATTTTAACTGAA 17
 RESULT 11
 A1571882/c 768 bp mRNA linear EST 12-MAY-1999
 LOCUS TO20F08.X1 NCI_CGAP_UT2 Homo sapiens cDNA clone IMAGE:2179623 3'
 DEFINITION Similar to gb:U14758 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED
 PROTEIN GA733-2 (HUMAN); mRNA sequence.
 ACCESSION A1571882
 VERSION A1571882.1 GI:4535256
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 768)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rtm1.nhl.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnl.gov/bdrrp/image/image.html
 Insert Length: 1567 Std Error: 0.00
 Seq primer: -40UP from Glibco
 High quality sequence stop: 414
 POLYA-No.

FEATURES

Location/Qualifiers
 1..768
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2179623"
 /clone_1lb="NCI-CGAP_Ut2"
 /tissue_type="moderately-differentiated endometrial
 adenocarcinoma, 3 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.85 kb. Life Technologies catalog #:
 11539-012"
 BASE COUNT 243 a 160 c 102 g 260 t 3 others
 ORIGIN

Query Match 67.2%; Score 451.4; DB 9%; Length 768;
 Best Local Similarity 99.8%; Pred. No. 2.6e-56;
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGATTAAGGAGATGGGAGATGATAGGACATCATGCTATATATTTGAA 215
 |
 DB 458 TGAGATAAAGGAGATGGGAGATGATAGGACATCATGCTATATATTTGAA 399
 QY 216 GATTATAGAGAAGGAGATGATAGGACATCATGCTATATATTTGAA 275
 |
 DB 398 GATTATAGAGAAGGAGATGATAGGACATCATGCTATATATTTGAA 339
 QY 276 AAGACATCTTTGAAGTGCATGATTTGTTAATCATCATATTTGTTAATGTA 335
 |
 DB 338 AAGACATCTTTGAAGTGCATGATTTGTTAATCATCATATTTGTTAATGTA 279
 QY 336 CCGTACTCAAAATATAGAGCTTGAACCTGCTTACCAATCTGAAATTTGACCACA 395
 |
 DB 278 CCGTACTCAAAATATAGAGCTTGAACCTGCTTACCAATCTGAAATTTGACCACA 219
 QY 396 AGTGTCTTATATGAGATCTATATGTAATCAATCCAGACCTGACATCGTTAAAT 455
 |
 DB 218 AGTGTCTTATATGAGATCTATATGTAATCAATCCAGACCTGACATCGTTAAAT 159
 QY 456 ATTATGTGTACATTCATGATGTCATTAATATGCTTCCACAGTAATCTGAAAA 515
 |
 DB 158 ATTATGTGTACATTCATGATGTCATTAATATGCTTCCACAGTAATCTGAAAA 99
 QY 516 CCGATTTGATGTAAGAGCTGCTTCTATTTACTGAGCTGCTGATACATACATTT 575
 |
 DB 98 CCGATTTGATGTAAGAGCTGCTTCTATTTACTGAGCTGCTGATACATACATTT 39
 QY 576 TATGAGCTATGAATAAATCAATTTTAACTGAA 608
 |
 DB 38 TATGAGCTATGAATAAATCAATTTTAACTGAA 6

RESULT 12
 AA502398/c 513 bp mRNA linear EST 19-AUG-1997
 LOCUS ne27h08.s1 NCI-CGAP_C03 Homo sapiens CDNA clone IMAGE:898623 3'
 DEFINITION similar to gb:U14758 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED
 PROTEIN GA733-2 (HUMAN);, mRNA sequence.
 ACCESSION AA502398
 VERSION AA502398.1 GI:2237365
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 513)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
 Ph.D.

cdna Library Preparation: M. Bento Soares, Ph.D.
 cdna Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnl.gov/bdrrp/image/image.html
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 352.
 Location/Qualifiers
 1..513

FEATURES

Location/Qualifiers
 1..513
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:898623"
 /clone_1lb="NCI-CGAP_C03"
 /sex="pooled"
 /tissue_type="colon"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from 12 pooled bulk tumor samples and primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT73 vector. Library went through one round of
 normalization."
 BASE COUNT 161 a 102 c 69 g 181 t
 ORIGIN

Query Match 67.0%; Score 450.4; DB 9%; Length 513;
 Best Local Similarity 99.8%; Pred. No. 4.6e-56;
 Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGATTAAGGAGATGGGAGATGATAGGACATCATGCTATATATTTGAA 215
 |
 DB 452 TGAGATAAAGGAGATGGGAGATGATAGGACATCATGCTATATATTTGAA 393
 QY 216 GATTATAGAGAAGGAGATGATAGGACATCATGCTATATATTTGAA 275
 |
 DB 392 GATTATAGAGAAGGAGATGATAGGACATCATGCTATATATTTGAA 333
 QY 276 AAGACATCTTTGAAGTGCATGATTTGTTAATCATCATATTTGTTAATGTA 335
 |
 DB 332 AAGACATCTTTGAAGTGCATGATTTGTTAATCATCATATTTGTTAATGTA 273
 QY 336 CCGTACTCAAAATATAGAGCTTGAACCTGCTTACCAATCTGAAATTTGACCACA 395
 |
 DB 272 CCGTACTCAAAATATAGAGCTTGAACCTGCTTACCAATCTGAAATTTGACCACA 213
 QY 396 AGTGTCTTATATGAGATCTATATGTAATCAATCCAGACCTGACATCGTTAAAT 455
 |
 DB 212 AGTGTCTTATATGAGATCTATATGTAATCAATCCAGACCTGACATCGTTAAAT 153
 QY 456 ATTATGTGTACATTCATGATGTCATTAATATGCTTCCACAGTAATCTGAAAA 515
 |
 DB 152 ATTATGTGTACATTCATGATGTCATTAATATGCTTCCACAGTAATCTGAAAA 93
 QY 516 CCGATTTGATGTAAGAGCTGCTTCTATTTACTGAGCTGCTGATACATACATTT 575
 |
 DB 92 CCGATTTGATGTAAGAGCTGCTTCTATTTACTGAGCTGCTGATACATACATTT 33
 QY 576 TATGAGCTATGAATAAATCAATTTTAACTGAA 607
 |
 DB 32 TATGAGCTATGAATAAATCAATTTTAACTGAA 1

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
RESULT 13 AI038566/c	AI038566	471 bp	mRNA	linear	EST 28-AUG-1998							
	0x34b10.s1 Soares-total_fetus_Nb2HF8_9w		Human sapiens cDNA clone									
	IMAGE:1658203	3'	similar to gb:U41758 MAJOR GASTROINTESTINAL TUMOR-ASSOCIATED PROTEIN GA73-2 (HUMAN);									
	AI038566											
	AI038566.1	GI:3277760										
	EST											
	human											
	Human sapiens											
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.											
	1 (bases 1 to 471)											
	NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap.										
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),											
	Tumor Gene Index											
	Unpublished (1997)											
	Contact: Robert Strausberg, Ph.D.											
	Email: cgapbs-remail.nih.gov											
	This clone is available royalty-free through LNL; contact the											
	IMAGE Consortium (infoimage.lnl.gov) for further information.											
	Insert Length: 1379	Std Error: 0.00										
	Seq primer: -40m13 fwd. ET from Amersham											
	High quality sequence stop: 367;											
	Location/Qualifiers											
	1..471											
	/organism="Homo sapiens"											
	/db_xref="taxon:9606"											
	/clone="IMAGE:1658203"											
	/clone_lib="Soares-total_fetus_Nb2HF8_9w"											
	/dev_stage="-8-9 weeks"											
	/lab_host="DH10B"											
	/note="Vector: pUT7D-Pac (Pharmacia) with a modified											
	polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA											
	was prepared from mRNA obtained from pooled 8-9 week											
	(total) fetus material with a Not I - oligo(dT) primer											
	TGTTACCACTTGAAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 371											
	Double-stranded cDNA was ligated to Eco RI adaptors											
	(Pharmacia), digested with Not I and cloned into the Not I											
	and Eco RI sites of the modified pUT73 vector. Library											
	went through one round of normalization, and was											
	constructed by Bento Soares and M. Fatima Bonaldo."											
	BASE COUNT	148 a	87 c	61 g	175 t							
	ORIGIN											
	Query Match	66.9%	Score 449.8	DB 9;	Length 471;							
	Best Local Similarity	99.6%	Pred. No. 5.9e-56;									
	Matches 451; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;							
Y	156 TCAGTAAAGGAGATGGGTGAGATGCATGAGGAGCTCAATGATATATAT											

Db	160	ATTTATGCTACATTCAATCATGTGGCATTAAATAATAGTCTCCACACTAAACTGGAAAA	101
Oy	516	CTGATTGTGATGAAGCTGCCCTTTCTAATTACTTGAGTCCTTGACATACACTACTTTT	575
Db	100	CGATTGTGATGTAAGAGCTGCCTTCTATTACTTGAGTCCTTGACATACACTACTTTT	41
Oy	576	TATGAGCTATGAATAAACATTTTAAACTGAA	608
Db	40	TATGAGCTATGAATAAACATTTTAAACTGAA	8
RESULT 14			
BMS10869/c			
LOCUS	BMS10869	481 bp	mRNA linear EST_15-REB-2007
DEFINITION	j144c04.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5633670		
ACCESSION	3' similar to SW:G732_HUMAN P16422 MAJOR GASTROINTESTINAL		
VERSION	TUMOR-ASSOCIATED PROTEIN GA733-2 PRECURSOR ; mRNA sequence.		
KEYWORDS	BMS10869		
SOURCE	BMS10869.1 GI:1868212		
ORGANISM	EST.		
REFERENCE	human.		
AUTHORS	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 481)		
JOURNAL	Melton, D., Brown, J.J., Kenty, G., Permutt, A., Lee, C., Kastner, K., Lemstra, L., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Page, D., Wyle, T., Martin, J., Blaisdell, A., Schmitt, A., Theisinger, B., Rutter, E., Ronko, J., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisshvili, R., Williams, T., Jackson, Y. and Bowers, Y. Endocrine Pancreas Consortium Unpublished (2000) Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@biohp.harvard.edu Library was constructed by Dr. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@m.wustl.edu) Seq primer: -40UP from Gibco.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..481		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="5633670"		
	/clone_lib="Human Insulinoma"		
	/tissue_type="Insulinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1: XhoI; Site: 2: EcoRI; Constructed with lambda ZAPI system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."		
BASE COUNT	146 a	88 c	63 g
ORIGIN	184 t		
Query Match	66.9%	Score 449.8	DB 13: Length 481;
Best Local Similarity	99.6%	Pred. No. 5.8e-56;	
Matches 451: Conservative	0:	Mismatches 2:	Indels 0: Gaps 0:

NAME/KEY: allele
LOCATION: 2934 : polymorphic base A or G
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947 : polymorphic base A or G
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802 : polymorphic base A or G
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062 : polymorphic base A or G
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088 : polymorphic base C or T
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109 : polymorphic base C or T
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170 : polymorphic base G or T
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903 : polymorphic base A or G
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019 : polymorphic base A or G
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141 : polymorphic base G or C
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183 : polymorphic base C or T
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338 : polymorphic base A or G
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375 : polymorphic base G or C
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429 : polymorphic base C or T
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467 : polymorphic base A or G
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484 : polymorphic base A or G
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534 : polymorphic base A or T
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611 : polymorphic base A or G
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668 : polymorphic base A or G
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608 : polymorphic base C or T
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658 : polymorphic base A or G
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703 : polymorphic base G or C
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777 : deletion of CTG
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785 : polymorphic base G or T
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele

LOCATION: 8926 : polymorphic base C or T
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171 : polymorphic base C or T
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429 : polymorphic base C or T
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341 : polymorphic base A or G
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492 : polymorphic base A or C
OTHER INFORMATION: 10-507-321 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 13524 : polymorphic base C or T
OTHER INFORMATION: 10-507-353 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13535 : polymorphic base C or T

Query Match 7.4% Score 50; DB 4; Length 20674;
Best Local Similarity 46.1% Pred. No. 0.095;
Matches 167; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 303 TTAGTTTAACTATATATTTGTAATAGTGAACCTGTAACCAATATACGAGCTTGA 362
DB 11103 TTAATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 11162
QY 363 AACTGCTTTACCAATCTGAAATTTGACCACAGTCTTATATATGACATCTAATGT 422
DB 11163 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAATTT 11222
QY 423 AAAATCCAGACTGTGACGCCATCGTAAATTTATTTAGTGAACATCAATGCTGC 482
DB 11223 TAAATTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 11282
QY 483 ATTAATATGCTTCCACAGTAAATCTGAAACAGTATTTGATGAAAGCTGCTTC 542
DB 11283 ACTTAATTTAATTTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 11342
QY 543 TATTTACTTGAGCTGTGACATACATACATTTTATGACCTATGAATTAATTAATTTAA 602
DB 11343 TTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAA 11402
QY 603 ACGAATTTCTTAACCTTGACATTTCAATTTCTTCTTTCTTTCTTTCTTTCTTTT 662
DB 11403 TTTTAATTTAATTTGCTCAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAA 11462
QY 663 TT 664
DB 11463 TT 11464

RESULT 4
US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2
```

```
Query Match
Best Local Similarity 54.3%; Score 49.6; DB 1; Length 5852;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```
QY 484 TTAATATGCTTCACAGTAATCTGAAAACGATTGGATTGAAAGCTGCTTCCT 543
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 5445 TTAATAAAAAAAAAAAAAAAAAAAAAATCAATATGTTTATGTTTAAAGATTTT 5504

QY 544 ATTACTGAGCTCTTGACATACACTTTTATGAGCTTGAAATTAACATTTTAA 603
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 5505 AATCTCGCAATGATTTTAAATATAAATCGATACATATTTTAAAAAACCCTTACA 5564

QY 604 CTGAATTTCTAATCTTGACATTTCAATTTCTCTCTTTCTTTCTTTCTTTT 663
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 5565 TTTTATTTTAAATTCACAAATTTATACATTTTATTTTATTTTATTTTATTT 5624
```

```
QY 664 TTTT 667
    ||||
Db 5625 TTTT 5628
```

```
RESULT 5
US-09-129-112-3/c
; Sequence 3, Application US/09129112
; Patent No. 6465716
; GENERAL INFORMATION:
; APPLICANT: Etzler, Marilyn E.
; APPLICANT: Murphy, Judith B.
; TITLE OF INVENTION: The Regents of the University of California
; FILE REFERENCE: 023070-079810US
; CURRENT APPLICATION NUMBER: US/09/129,112
; CURRENT FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 08/907,226
; FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6265
; TYPE: DNA
```

```
ORGANISM: Dolichos biflorus
FEATURE:
OTHER INFORMATION: genomic sequence of NBP46 (DB46)
NAME/KEY: exon
LOCATION: (633)..(944)
NAME/KEY: intron
LOCATION: (945)..(1022)
NAME/KEY: exon
LOCATION: (1023)..(1151)
NAME/KEY: intron
LOCATION: (1152)..(1559)
NAME/KEY: exon
LOCATION: (1560)..(1616)
NAME/KEY: intron
LOCATION: (1617)..(1697)
NAME/KEY: exon
LOCATION: (1698)..(1790)
US-09-129-112-3
```

```
Query Match
Best Local Similarity 47.2%; Score 49.2; DB 4; Length 6265;
Matches 150; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
```

```
QY 351 TAAGCAGCTTGAAACGCTTTACCAATCTGAAATTTGACACAGCTGCTTATATG 410
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3720 TTAAGAGTAATAAATTGACAGAAAAAATACATATGATATTAATTAATATAT 3661

QY 411 CAGATCTAATGTAAATCCAGAACTTGACCTCCATCGTTAAATTTATTTATGTA 470
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3660 AACATTAATAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3601

QY 471 TCAATGCTGTCATTAATATGCTTCCACAGTAATAATCTGAAAACGATTGGATTG 530
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3600 ATATTTATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3541

QY 531 AAGTCGCTTCTATTTACTTGAGCTGTTGACATACATTTTATTTATGAGCTATGA 550
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3540 TAAATTTAAATAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 3481

QY 591 AAACATTTTAACTGAATTTCTTAATTTGACATTTCAATTTCTCTCTTTCTTT 650
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3480 ATTAATAATAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTAT 3421

QY 651 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 668
    ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 3420 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3403
```

```
RESULT 6
US-08-213-419B-3/c
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; FILE REFERENCE: JTI-002CNCPE
; CURRENT APPLICATION NUMBER: US/08/213,419B
; CURRENT FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (2598)..(3404)
; NAME/KEY: CDS
```

```
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
US-08-213-419B-3
```

Query Match	7.28;	Score: 48.6;	DB 4;	Length 6124;
Best Local Similarity	64.98;	Pred. No. 0.16;		
Matches	72;	Conservative	0;	Mismatches 39;
			Indels	0;
			Gaps	0

[illegible]

```

RESULT 7
US-08-973-462-1
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERTS, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973.462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
; US-08-973-462-1

```

Query Match	6.9%	Score 46.4	DB 4	Length 6152
Best Local Similarity	48.0%	Pred. NO.0.42		
Matches 169	Conservative 0	Mismatches 176	Indels 7	Gaps 1
QY 316	ATATATTGTATAGTGAACCTGTACTACCAAAATATAGCAGCTGGAACCTGGCTTACC	375		
Db 5720	ATATATATATATATATATGATATCTTTTACAAAATTTTAAATTTTATATATAT	5779		
QY 376	AATCTTGAATTTGACCAAGTGCTTATATATATGACAGATCTAATGTAAATCCAGACT	435		
Db 5780	AATATTATATATTTTCCATATATATTTTATTTTCAATATTTTATTTTAAATTAATGA	5839		
QY 436	TGGACATCATGTTAAATTTATTTATGTTACATCTCAATGTGTCATTAATATAGCT	495		
Db 5840	TTTTTTACAGATGTATGTGTTTTATTAATTAATATATAGATTTCTGTAGAAA-----CT	5892		
QY 496	CCACAGTAAATATCTGAAAACATGATTTGATTTGATTTGAAGCGCTTCTCTTACTTGAGT	555		
Db 5893	GTAATATTTTCATACGATATATAGTAATATTAATTTATTTGTGTTTTTAAAAATTTATAT	5952		
QY 556	CTTGACATACATACCTTTTTATATGAGCATATGAATAAAAACATTTAACTGATTTCTTAA	615		
Db 5953	ATAGAAAGATAAAATTTAGCTATATT	6012		
QY 616	ACTTGTACATTTCAAAATTTCTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	667		
Db 6013	TGCTGTATATGCAAAATTAAGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATAT	6064		

RESULT 8
US-08-451-405A-2
; Sequence 2, Application US/08451405A

Patent No. 5736358
GENERAL INFORMATION:
APPLICANT: FASEL, NICOLAS JOSEPH
APPLICANT: REYMOND, CHRISTOPHE DOMINIQUE
TITLE OF INVENTION: DICTYOSHELID EXPRESSION VECTOR AND
TITLE OF INVENTION: METHOD FOR EXPRESSING A DESIRED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE WEBB LAW FIRM
STREET: 700 KOPPEERS BUILDING, 436 SEVENTH AVENUE
CITY: PITTSBURGH
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 15219-1818
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 FLOPPY DISK
COMPUTER: Midwest Micro 486-50
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,405A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,273
FILING DATE: 15-JAN-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-08-451-405A-2

Query Match	6.9%;	Score 46.2;	DB 1;	Length 731;
Best Local Similarity	44.3%;	Pred. No. 0.36;		
Matches 276;	Conservative	0;	Mismatches 343;	Indels 4;
			Gaps	2

Oy	9	TTTTTCCTTTTCTTTTAAATTAAGCTATAAAATGTGGAAAAAATTATCTGTGCTCT	68
Dd	83	TTATTTTATTTTATTTTATTTTAAAAAAATAAAAATTAAGATAAAATATTTCTATTGAG	142
Oy	69	TTAATTTCTATTTTATTTAATTAAGTATTTTCAGAAATGAACAAGATTTGAAAAATTAATTA	128
Dd	143	GAGTTTATTTTATTTGATTTTAAAAATTAATTAATTAACCTAGGAACCTAAAAATAGATTTGTGA	202
Oy	129	GAAATTTTCTGTGCTTTTCCGCTTTCAGATTAAGAGATGGCTGATGACCTATAGGA	188
Dd	203	CGGTTATGATTAAGAAAAATTTCTAAAAAAAATTTTCAGTAAATTTTGTGATTGGAAACAC	265
Oy	189	ACTCAATGCATTAAGTATTTAATTTTGAAGATTTTGAAGAAAGAAAGAAATGCAAC	248
Dd	263	AACCA-----AAATTA	319
Oy	249	AAATTAACAATGTGTGCGCTGGGACGACAGACATCTTGAAGGTCATGAGCTTTCTACTT	308
Dd	320	AAATCAAAAAAAAAAAGTATTTAAAGAAATTTTAAAAATTTATTAATATCTTTTAAT	375
Oy	309	TAACTCATATATTTTGTATATAGTGAACCTGTACTCAAAATATTAAGCAGCTTGAACGTG	368
Dd	380	GTGCAAAACACACTTTTAACACACTCTATTTATCTTACAAAGCTTTAAATTTTAATTTT	433
Oy	369	CTTTACCAATCTTGAATTTGACACACAGTGTCTTAATATATGACAGATCTAATGTAAATC	422
Dd	440	TTTATTTATTAATTAATTTTAAAAATTAATTTTAAAAATTTTAAATTTTAAATTTT	493
Oy	429	CAGACTTGGACTCATGCTGTTAAAAATTTTATGTGTAACTCAAAATGTGCGATTA	488
Dd	500	TTTACATCAACCCCTTAACTCAAAACAATACATTTATTTTATTTATTTTAATATA	555
Oy	489	TATGCTTCCACAGTAAATCTGAAAAAAGTATTTGTGATGAAGCTGCCCTTCTATTTA	548
Dd	560	T-CATTTGAAATATAAAATATTTTCTAATGTGTAGTATATAATTTCTTTTAAATATA	618

US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marile
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296, 95238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note="RP4 origin of DNA transfer (orig) from
; OTHER INFORMATION: plasmid RP4"
; US-09-056-075-1

Query Match 6.6%; Score 44.2; DB 2; Length 6243;
Best Local Similarity 58.9%; Pred. No. 1.1;
Matches 76; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 539 TTTCTATTACTGAGCTGTGATACATACATCTTTTATGAGCTAGTAATAACATT 598
DB 1464 TCTTTCTCTCAAGATATATATAATAAATTTTTCACAACTTAATAAATAA 1405
QY 599 TTAACGTAATTTCTTAACTTGACATTTCAATTTCTTCTTTCTTTCTTTT 658
DB 1404 TTTTATATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1345
QY 659 TTTTATTTT 667
DB 1344 TTTTATTTT 1336

RESULT 12
US-08-446-855A-1/c
; Sequence 1, Application US/08446855A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria V
; APPLICANT: O'Sullivan, William J
; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl

; TITLE OF INVENTION: phosphate synthetase II
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,855A
; FILING DATE: 06-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 47-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic
; US-08-446-855A-1

Query Match 6.5%; Score 43.6; DB 2; Length 8920;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 109; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 448 TTAATAATTTATTTATGTCATCAATGTCATTAATTAATGCTCCACAGTAAT 507
DB 834 TAAACGTTTAAAGACCATATGATTCAGAGATATCCCAATATATATATATAT 775
QY 508 CTGAACACGATTTGATGATGAAGAGTCCTTTCTATTTTACTTGATCTGTACAT 567
DB 774 TTTATATATATATCATATATATTTTTCCTTTCTTTTATTTTATATATAT 715
QY 568 TACTTTTATGAGCTATGAATAATAACATTTTAACTGATTTCTTACTTGACATT 627
DB 714 TTTATGTTTAAATAATTTATTAATTTACATATACAAAGTTCAATATGTAATTT 655
QY 628 CAAATTTCTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTT 665
DB 654 TTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 617

RESULT 13
US-09-150-741-1/c
; Sequence 1, Application US/09150741
; Patent No. 618396
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; TITLE OF INVENTION: Synthetase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06

NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 8920
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-09-150-741-1

Query Match
Best Local Similarity 50.0%; Score 43.6; DB 4; Length 8920;
Pred. No. 1.6;
Matches 109; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 448 TTTAAATTTTATGCTGATTCATTAAGTGTGATTAATATGCTTCACAGTAAAT 507
DB 834 TAAACTGTTTAAAGCTTATGATTCAGAGATATCCCAATATATATATATATA 775
QY 508 CTGAAACATGATTTGTGATTTGAACCTGCTTATTTACTTGATCTTGACATACA 567
DB 774 TTTATATATATATCTATATATATTTTCCCATTTTCTTTTATATACATTTATA 715
QY 568 TACTTTTATGAGCTATGAAATAAACATTTTAACTGATTTTGACATTT 627
DB 714 TTTATGTTTAAATATTTTAAATTTTACATATACAGTTTTCATATGTAATTTT 655
QY 628 CAAATTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 665
DB 654 TTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 617

RESULT 14
US-09-426-290-1
Sequence 1, Application US/09426290
Patent No. 6410712
GENERAL INFORMATION:
APPLICANT: Berglund Ran Oiafsdottir
TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
FILE REFERENCE: 2345.2001-000
CURRENT APPLICATION NUMBER: US/09/426,290
CURRENT FILING DATE: 1999-10-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 168575
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (21181)...(21403)
NAME/KEY: CDS
LOCATION: (95252)...(95430)
NAME/KEY: CDS
LOCATION: (101753)...(101996)
NAME/KEY: CDS
LOCATION: (110324)...(110439)
NAME/KEY: CDS
LOCATION: (124058)...(124278)
NAME/KEY: CDS
LOCATION: (127009)...(127130)
NAME/KEY: CDS
LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match
Best Local Similarity 48.0%; Score 43.4; DB 4; Length 168575;
Pred. No. 2.4;
Matches 154; Conservative 0; Mismatches 166; Indels 1; Gaps 1;

QY 347 AATATAGCAGCTTAAACCTGCTTACCAATCTTGAATTTGACCACAAAGTGTCTTATA 406
DB 120800 AATATATATATATAGAGATTTTATATCTCTGCAATGTTTATATATTTTCAAGTATA 120859
QY 407 TATGAGATCTATATGTAATATCCAGACTGACATCGTTAAATATATTTATGTGTA 466

DB 120860 TATATATATATATGAAATATATATCTTATACATATATTTTATGAAATATATACTTA 120919
QY 467 ACATTCAAATGTGCTGATTAATATGCTTCACAGTAAATCTGAAACATGATTTGTGA 526
DB 120920 TATATATATTTTATGAAATATATATACTTATATATATTTTATGAAATATATACTTA 120979
QY 527 TTGAAAGCTGCTTCTTATTTACTTGAAGCTTGTGATACATACATCTTTTATGAGCTATG 586
DB 120980 TATATATATTTTATGAAATATA-TATACTTATATATATTTTATGAAATGATGCTTAA 121038
QY 587 AATTAACATTTTAAACGAAATTTCTTAACTTGACATTTTCAATTTCTTCTTTT 646
DB 121039 AACACATTTCTTTGATTTGAAACCTTTTAAATTTTATTTTCCATTTATTTGCTAGTATGT 121098
QY 647 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 667
DB 121099 AGAAGTATATGATTTTCTTTT 121119

RESULT 15
US-08-487-826B-13/c
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESSES:
ADDRESS: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match
Best Local Similarity 43.9%; Score 43; DB 2; Length 19124;
Pred. No. 2.2;
Matches 293; Conservative 0; Mismatches 360; Indels 14; Gaps 2;
QY 2 AATATACCTTTTCTTTTATTTAAATGCTTATATAAGTGGAAAAATATATCTTG 61

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	672	100.0	672	10	US-09-919-407-32	Sequence 32, App
2	672	100.0	672	10	US-09-880-197-2421	Sequence 2421, App
3	461.6	68.7	621	10	US-09-777-564-191	Sequence 191, App
4	451.4	67.2	1597	10	US-09-925-300-656	Sequence 656, App
5	449.4	66.9	1538	10	US-09-922-217-1106	Sequence 1106, App
6	443.4	66.0	637	10	US-09-920-300A-14	Sequence 14, App
7	443.4	66.0	637	12	US-10-033-528-14	Sequence 14, App
8	442.4	65.8	481	10	US-09-920-300A-1559	Sequence 1559, App
9	442.4	65.8	481	12	US-10-033-528-1559	Sequence 1559, App
10	439.4	65.4	563	10	US-09-998-598-235	Sequence 235, App
11	439	65.3	439	10	US-09-777-564-1254	Sequence 1254, App
12	433.4	64.5	534	10	US-09-998-598-290	Sequence 290, App
13	430.2	64.0	566	10	US-09-998-598-1423	Sequence 1423, App
14	415.4	61.8	425	10	US-09-922-217-591	Sequence 591, App
15	415.4	61.8	425	10	US-09-833-263-591	Sequence 591, App
16	335.8	50.0	507	10	US-09-777-564-1541	Sequence 1541, App
17	333.8	49.7	555	10	US-09-920-300A-7	Sequence 7, App
18	333.8	49.7	555	12	US-10-033-528-7	Sequence 7, App
19	332.2	49.4	478	10	US-09-998-598-338	Sequence 338, App

C	20	326	48.5	326	12	US-09-920-100A-445	Sequence 445, App
C	21	326	48.5	326	10	US-10-033-828-445	Sequence 445, App
C	22	269	40.0	269	10	US-09-922-217-137	Sequence 137, App
C	23	269	40.0	269	10	US-09-833-763-137	Sequence 137, App
C	24	255.4	38.0	257	10	US-09-998-998-92	Sequence 92, Appl
C	25	224.4	33.4	261	10	US-09-998-998-254.3	Sequence 254.3, App
C	26	220	32.7	220	9	US-10-046-935-1076	Sequence 1076, Ap
C	27	220	32.7	220	9	US-10-046-935-1076	Sequence 1952, Ap
C	28	219	32.6	220	9	US-10-046-935-960	Sequence 960, App
C	29	205.8	30.6	209	10	US-09-998-998-2027	Sequence 2027, App
C	30	201.4	30.0	217	10	US-09-998-998-347	Sequence 347, App
C	31	198	29.5	198	9	US-10-046-935-2188	Sequence 2188, Ap
C	32	185.4	27.6	357	10	US-09-998-998-2544	Sequence 2544, Ap
C	33	183.4	27.3	235	9	US-10-046-935-421	Sequence 421, App
C	34	183.4	27.3	554	10	US-09-922-217-413	Sequence 413, App
C	35	183.4	27.3	554	10	US-09-833-763-413	Sequence 413, App
C	36	182.4	27.1	202	9	US-10-046-935-1624	Sequence 1624, Ap
C	37	178	26.5	178	10	US-09-920-100A-1486	Sequence 1486, Ap
C	38	178	26.5	178	12	US-10-033-928-1486	Sequence 1486, Ap
C	39	171.4	25.5	459	9	US-10-046-935-20	Sequence 20, Appl
C	40	148	22.0	373	9	US-10-042-125A-38	Sequence 38, Appl
C	41	145	21.6	160	10	US-09-777-864-325	Sequence 325, App
C	42	107.4	16.0	231	10	US-09-876-889-114	Sequence 114, App
C	43	104.4	15.5	106	10	US-09-998-998-169	Sequence 169, App
C	44	52.4	7.8	411	10	US-09-960-552-14521	Sequence 14521, App
C	45	51.4	7.6	424	10	US-09-960-352-11218	Sequence 11218, A

ALIGNMENTS

RESULT 1

```

: Sequence 32, Application US/09919497
: Patient No. US200201066521
:
: GENERAL INFORMATION:
:
: APPLICANT: Mutter, George L.
: TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
: FILE REFERENCE: B0801,7225
: CURRENT APPLICATION NUMBER: US/09/919,497
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/221,735
: PRIOR FILING DATE: 2000-07-31
: NUMBER OF SEQ ID NOS: 100
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 32
: LENGTH: 672
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
: US-09-919-497-32

```

Query Match	Score	DB	Length
100.0%	672	10	672

Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATATACTTTTCTTTTATTAAATAAGCTTATAAATGTGGGAAAAAATTATCTT 60

Accession	Gene	Position	Sequence	Position
Db	1	AAATTAACCTTTTCTTTTTTATTTAAATAAGCTTTTAAATGTGGCAAAAAATATCTT	60	
Qy	61	GTATCCCTTAATTCATTTTATTTAATCTATTTTCAGATGAAACAAAAGATTGAAA	120	
Db	61	GTGTTCCTTAATTCATTTTATTTAATACATATTTTCAGATGAACAAAAGATTGAAA	120	
Qy	121	ATATATTAGAAATTTTTTCTGTGCTTTTCCCTTTTCAGATTAAGGACATGGGTGAGATG	180	
Db	121	ATATATTAGAAATTTTTTCTGTGCTTTTCTGTTTCAGATTAAGGACATGGGTGAGATG	180	
Qy	181	CATAGGGAACCTCAATGCATACTATATTAATTTGAAGATTATAGAAAGAGAAATATACAA	240	
Db	181	CATAGGGAACCTCAATGCATACTATATTAATTTGAAGATTATAGAAAGAGAAATATACAA	240	
Qy	241	ATGACACAAATTTCAAAATGTGTGTGGTGGGACGAACACATCTTTTAAAGTCAATGAGTT	300	
Db	241	ATGACACAAATTTCAAAATGTGTGTGGTGGGACGAACACATCTTTTAAAGTCAATGAGTT	300	

Db 241 ATGACACAAATTTACAAATGTGTGCGTGGAGCAAGACATCTTTGAAGTCATGAGTT 300
Qy 301 TGTGATTACATCATATATTTGTATAGTGAAACCTGTACTCAAAATATAGAAGCTT 360
Db 301 TGTGATTACATCATATATTTGTATAGTGAAACCTGTACTCAAAATATAGAAGCTT 360
Qy 361 GAACTGGCTTTACCAATCTTGAATTTTGACCAAGTCTTTATATATGACAGATCTAAT 420
Db 361 GAACTGGCTTTACCAATCTTGAATTTTGACCAAGTCTTTATATATGACAGATCTAAT 420
Qy 421 GTTAAATCCAGAACTGGACCTCATCGTTAAATTTATATGTATGTAACATTCATATGT 480
Db 421 GTTAAATCCAGAACTGGACCTCATCGTTAAATTTATATGTATGTAACATTCATATGT 480
Qy 481 GCATTAAATATGCTTCACAGTAAATCTGAAAACTGATTTGTGATTAAGAGCTGCTT 540
Db 481 GCATTAAATATGCTTCACAGTAAATCTGAAAACTGATTTGTGATTAAGAGCTGCTT 540
Qy 541 TCTATTACTGAGCTCTGTACATACATACATCTTTTATGAGCTATGAAATTAACAATTTT 600
Db 541 TCTATTACTGAGCTCTGTACATACATACATCTTTTATGAGCTATGAAATTAACAATTTT 600
Qy 601 AAACGAAATTTCTTACCTTGACATTTCAAAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
Db 601 AAACGAAATTTCTTACCTTGACATTTCAAAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
Qy 661 TTTTCTTTGAGA 672
Db 661 TTTTCTTTGAGA 672

RESULT 2

US-09-880-107-2421
; Sequence 2421, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles In Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2421
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M93036
US-09-880-107-2421

Query Match 100.0%; Score 672; DB 10; Length 672;
Best Local Similarity 100.0%; Pred. No. 5e-105;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATATACCTTTCTTTTATTTTAAATAGCTATTAATGCGGAAAAATATATCTT 60
Db 1 AATATACCTTTCTTTTATTTTAAATAGCTATTAATGCGGAAAAATATATCTT 60
Qy 61 GGTGTCCTTTAATTCATTTTATTTAATATGATTTTGAAGTGAATTTGAAAA 120
Db 61 GGTGTCCTTTAATTCATTTTATTTAATATGATTTTGAAGTGAATTTGAAAA 120
Qy 121 ATTATTTAATTTTCTGCTCTTTCCCTTTCAATATAAGAGATGGTGAGTG 180
Db 121 ATTATTTAATTTTCTGCTCTTTCCCTTTCAATATAAGAGATGGTGAGTG 180

Qy 181 CATAGGAGCTCAATCATATATATTTGAAGATTATAGAAAGGAAATAGCAA 240
Db 181 CATAGGAGCTCAATCATATATATTTGAAGATTATAGAAAGGAAATAGCAA 240
Qy 241 ATGACACAAATTTACAAATGTGTGCGTGGAGCAAGACATCTTTGAAGTCATGAGTT 300
Db 241 ATGACACAAATTTACAAATGTGTGCGTGGAGCAAGACATCTTTGAAGTCATGAGTT 300
Qy 301 TGTGATTACATCATATATTTGTATAGTGAAACCTGTACTCAAAATATAGAAGCTT 360
Db 301 TGTGATTACATCATATATTTGTATAGTGAAACCTGTACTCAAAATATAGAAGCTT 360
Qy 361 GAACTGGCTTTACCAATCTTGAATTTTGACCAAGTCTTTATATATGACAGATCTAAT 420
Db 361 GAACTGGCTTTACCAATCTTGAATTTTGACCAAGTCTTTATATATGACAGATCTAAT 420
Qy 421 GTTAAATCCAGAACTGGACCTCATCGTTAAATTTATATGTATGTAACATTCATATGT 480
Db 421 GTTAAATCCAGAACTGGACCTCATCGTTAAATTTATATGTATGTAACATTCATATGT 480
Qy 481 GCATTAAATATGCTTCACAGTAAATCTGAAAACTGATTTGTGATTAAGAGCTGCTT 540
Db 481 GCATTAAATATGCTTCACAGTAAATCTGAAAACTGATTTGTGATTAAGAGCTGCTT 540
Qy 541 TCTATTACTGAGCTCTGTACATACATACATCTTTTATGAGCTATGAAATTAACAATTTT 600
Db 541 TCTATTACTGAGCTCTGTACATACATACATCTTTTATGAGCTATGAAATTAACAATTTT 600
Qy 601 AAACGAAATTTCTTACCTTGACATTTCAAAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
Db 601 AAACGAAATTTCTTACCTTGACATTTCAAAATTTCTTCTTTCTTTCTTTCTTTCTTTT 660
Qy 661 TTTTCTTTGAGA 672
Db 661 TTTTCTTTGAGA 672

RESULT 3

US-09-777-564-191/c
; Sequence 191, Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 191
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(621)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-564-191

Query Match 68.7%; Score 461.6; DB 10; Length 621;
Best Local Similarity 91.3%; Pred. No. 1.1e-69;
Matches 525; Conservative 0; Mismatches 44; Indels 6; Gaps 4;

Qy 26 AATAAGCTTTAATAATGCGGAAAAATATCTGTCTCTTAAATTCATTTTAT 85
Db 569 AATAAGTWTATTAATGGGAAAAATTTCTGTG--TTCNTTAATTCATTTTAT 513
Qy 86 TAAATATATTTTCAAGATGAACAAAGATGMAAAATTTATAGATTTTCTGTGCT 145
Db 512 TAANGCTATTTTCAAGATG-ACNAAGAGATTGAAAAATTTATTTAG-AATTTTCTGTGCT 455

	Query Match	Similarity	Best Local	Score	DB	Length
	Matches	452: Conservative	99.8%: Pred	51.4: 6.4e-68	1597:	
QY	156	TCAGATAAAGGAGATGGGTGAGATGCATATGGGAACTCAATGCATACTATATATTTGAA	1147			
DB	1088	TGAGATAAAGGAGATGGGTGAGATGCATATGGGAACTCAATGCATACTATATATTTGAA	1147			
QY	216	GATTATAGAAAGGGAATAGCAATTCACAAATTTCAAAATGTTGTTGCTGGGAGC	275			
DB	1148	GATTATAGAAAGGGAATAGCAATTCACAAATTTCAAAATGTTGTTGCTGGGAGC	1207			
QY	276	AAGACATCTTTGAAGGTCTAGATTTGTTAGTTAACTCATATATTTTAAATAGGAAA	335			
DB	1208	AAGACATCTTTGAAGGTCTAGATTTGTTAGTTAACTCATATATTTTAAATAGGAAA	1267			

Query Match	Similarity	Score	44.9	4	DB	10	Length	1528
Best Local	Similarity	99.8%	Pred.	No.	1.4e-67			
Matches	450	Conservative	0	Mismatches	1	Indels	0	Gaps
QY	156	TCAGATAAAGGAGATGGTGAGATGCATATGGGAAACCAATGCATGCATATATATATTTGAA	215					
Db	1078	TCAGATAAAGGAGATGGTGAGATGCATATGGGAAACCAATGCATGCATATATATATTTGAA	1137					
QY	216	GATTATAGAAGGAAGGAATAGCAAAATGCACAAATTACAAATGCTGTGTGGTGGGACG	275					
Db	1138	GATTATAGAAGGAAGGAATAGCAAAATGCACAAATTCAAATATGCTGTGGTGGGACG	1197					
QY	276	AAGACATCTTTGAAGGTCATGATGTTGGTTAGTTTAACATCATATATTTTGAATAGTAAA	335					
Db	1198	AAGACATCTTTGAAGGTCATGATGTTGGTTAGTTTAACATCATATATTTTGAATAGTAAA	1257					
QY	336	CCGTGATCTCAAAATATATACACAGCTTGAAACTGGCTTTACCAATCTTGAATTTGACCACA	395					
Db	1258	CCGTGATCTCAAAATATATACACAGCTTGAAACTGGCTTTACCAATCTTGAATTTGACCACA	1317					
QY	396	AGTGTCTTATATATGCAGATCTTAATGTAAATTCAGAACTTGGACATTCATCGTTAAATTT	455					

Db 1318 AGTGTCTATATATGAGATCTAATGTAAATCCAGAACTTGACATCGTTAAATTT 1377
Qy 456 ATTATAGTAACTTCAATGATGTCATTAATATGCTCCAGAGTAAATCTGAAAA 515
Db 1378 ATTATAGTAACTTCAATGATGTCATTAATATGCTCCAGAGTAAATCTGAAAA 1437
Qy 516 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 575
Db 1438 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 1497
Qy 576 TATGAGCTATGAATAAATCAATTTAACTG 606
Db 1498 TATGAGCTATGAATAAATCAATTTAACTG 1528

RESULT 6

US-09-920-300A-14/C
; Sequence 14, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-14

Query Match 66.0%; Score 443.4; DB 10; Length 637;
Best Local Similarity 99.8%; Pred. No. 1.3e-66;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 156 TCAGATAAGAGATGGGTGAGATGCAATGAGAACTCATACATATATTAATTTGAA 215
Db 445 TGAGATAAGAGATGGGTGAGATGCAATGAGAACTCATACATATATTAATTTGAA 386
Qy 216 GATTATAGAAGAGGAAATAGCAAAATGACACAAATTAACAATGTGTGCGGGGAGC 275
Db 385 GATTATAGAAGAGGAAATAGCAAAATGACACAAATTAACAATGTGTGCGGGGAGC 326
Qy 276 AAGACATCTTTGAAGGTCATGAGTTTGTATTAACTCATATATTTGTAATAGTGAAA 335
Db 325 AAGACATCTTTGAAGGTCATGAGTTTGTATTAACTCATATATTTGTAATAGTGAAA 266
Qy 336 CCTGACTCAAAATATTAAGACGCTTGAACCTGGCTTTACCAATCTTGAAATTTGACACA 395
Db 265 CCTGACTCAAAATATTAAGACGCTTGAACCTGGCTTTACCAATCTTGAAATTTGACACA 206
Qy 396 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGAGCTCATGCTTAAATTT 455
Db 205 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGAGCTCATGCTTAAATTT 146
Qy 456 ATTATATGTTACATTTCAATGTGTGCAATTAATATGCTTCCACAGTAATCTGAAAAA 515
Db 145 ATTATATGTTACATTTCAATGTGTGCAATTAATATGCTTCCACAGTAATCTGAAAAA 86
Qy 516 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 575
Db 85 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 26
Qy 576 TATGAGCTATGAATAAATCAATTT 600
Db 25 TATGAGCTATGAATAAATCAATTT 1

RESULT 7

US-10-033-528-14/C
; Sequence 14, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033.528
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-14

Query Match 66.0%; Score 443.4; DB 12; Length 637;
Best Local Similarity 99.8%; Pred. No. 1.3e-66;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 156 TCAGATAAGAGATGGGTGAGATGCAATGAGAACTCATACATATATTAATTTGAA 215
Db 445 TGAGATAAGAGATGGGTGAGATGCAATGAGAACTCATACATATATTAATTTGAA 386
Qy 216 GATTATAGAAGAGGAAATAGCAAAATGACACAAATTAACAATGTGTGCGGGGAGC 275
Db 385 GATTATAGAAGAGGAAATAGCAAAATGACACAAATTAACAATGTGTGCGGGGAGC 326
Qy 276 AAGACATCTTTGAAGGTCATGAGTTTGTATTAACTCATATATTTGTAATAGTGAAA 335
Db 325 AAGACATCTTTGAAGGTCATGAGTTTGTATTAACTCATATATTTGTAATAGTGAAA 266
Qy 336 CCTGACTCAAAATATTAAGACGCTTGAACCTGGCTTTACCAATCTTGAAATTTGACACA 395
Db 265 CCTGACTCAAAATATTAAGACGCTTGAACCTGGCTTTACCAATCTTGAAATTTGACACA 206
Qy 396 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGAGCTCATGCTTAAATTT 455
Db 205 AGTGTCTATATATGACAGATCTAATGTAAATCCAGAACTGAGCTCATGCTTAAATTT 146
Qy 456 ATTATATGTTACATTTCAATGTGTGCAATTAATATGCTTCCACAGTAATCTGAAAAA 515
Db 145 ATTATATGTTACATTTCAATGTGTGCAATTAATATGCTTCCACAGTAATCTGAAAAA 86
Qy 516 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 575
Db 85 CTGATTTGTGATGAAAGCTGCTTCTTAATTTACTTGAGCTCTGTACATCACTTTT 26
Qy 576 TATGAGCTATGAATAAATCAATTT 600
Db 25 TATGAGCTATGAATAAATCAATTT 1

RESULT 8

US-09-920-300A-1559/C
; Sequence 1559, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A

```

? CURRENT FILING DATE: 2001-07-31
? NUMBER OF SEO ID NOS: 1789
? SOFTWARE: PstSsd for Windows Version 4.0
? SEO ID NO 1559
? LENGTH: 481
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 354
? OTHER INFORMATION: n = A,T,C or G
? OS-09-920-300A-1559

```

Query Match	95.88;	Score 442.4;	DB 10;	Length 481;
Best Local Similarity	69.68;	Pred. No. 1.8e-66;		
Matches 443;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

OY	156	TCACATTAAGGAGATGGGTGAGATGACATAGGAACTCAATGCAATCAATATATTTGAA	215
Db	445	TGAGATTAAAGGAGATGGGTGAGATGACATAGGAACTCAATGCAATCAATATATTTGAA	386
OY	216	GATTATAGAGAGAGGAATAGCAAATGGACACAATTTACAATGTGTGCGTGGACG	275
Db	385	GATTATAGAGAGAGGAATAGCAAATGGACACNCAAATTTACAAATGTGTGCGCGGGACG	328
OY	276	AAGACATGTTGAAGGCTATAGATTTGTTACTTTACATCATATATTTGTAATAGTGAAA	333
Db	335	AAGACATGTTGAAGGCTATAGATTTGTTACTTTACATCATATATTTGTAATAGTGAAA	266
OY	336	CGTGACTCAAAATATAGACAGCTTGAACACGCGCTTACCAATCTTGAATTTGACCACA	395
Db	265	CGTGACTCAAAATATAGACAGCTTGAACACGCGCTTACCAATCTTGAATTTGACCACA	206
OY	336	AGTGTCTATATATAGCAATCTAATGTAAATCCAGAACTTGGACCTCCATCGTTAAATT	455
Db	205	AGTGTCTATATATAGCAATCTAATGTAAATCCAGAACTTGGACCTCCATCGTTAAATT	146
OY	456	ATTATATGCTAACTTCAATGTGTGACTTAATATGCTCCACAGTAAATCTGAAAAA	515
Db	145	ATTATATGCTAACTTCAAAATGTGTGACTTAAATATGTCTCCACAGTAAATCTGAAAAA	86
OY	516	CTGATTTTGATTGAAGAGCTGCCTTTCTATTTACTTGTAGTCCTTGATACATACCTTTT	578
Db	85	CTGATTTTGATTGAAGAGCTGCCTTTCTATTTACTTGTAGTCCTTGATACATACCTTTT	26
OY	576	TATGAGCTATGAATAAACAATTT 600	
Db	25	TATGAGCTATGAATAAACAATTTT 1	

```

RESULT 9
US-10-033-528-1559/c
Sequence 1559, Application US/10033528
Patent No. US20020131971A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jianshun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF COLON CANCER
FILE REFERENCE: 210121.547C1
CURRENT APPLICATION NUMBER: US/10/033,528
CURRENT FILING DATE: 2001-12-26
NUMBER OF SEQ ID NOS: 1896
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1559
LENGTH: 481
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 354

```

OTHER INFORMATION: n - A, T, C or G
US-10-033-528-1559

Query Match	65.8%	Score 442.4	DB 12	Length 481
Best Local Similarity	99.6%	Pred. No. 1.8e-66		
Matches 443, Conservative	0	Mismatches 2	Indels 0	Gaps 0

OY	156	TCAGCTAAAGAGATGGGTGAGATGCATAGGGAACCTCAATGCATATACATATATATTTGAA	215
Db	445	TGAGATTAAGSAGATGGTGTAGATGCATAGGGAACCTCAATGCATATATATATTTGAA	386
OY	216	GATTATAGAGAAGGGAATATGCAATAGCACACAAATTTACAATGTGTGTGGCTGGAGC	275
Db	385	GATTATAGAGAAGGGAATATGCAAAATGAGCNCMAATTTACAATGTGTGTGGCTGGAGC	326
OY	276	AAGCATCTTTGAAGGCATAGCTTTGTAGTTAACAATATATTTGTATATAGTGAA	335
Db	325	AAGCATCTTTGAAGGCATAGCTTTGTAGTTAACAATATATTTGTATATAGTGAA	286
OY	336	CCTGTACTCAAAATATATAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCACA	395
Db	265	CCTGTACTCAAAATATATAGCAGCTTGAACCTGGCTTTACCAATCTTGAATTTGACCACA	206
OY	366	AGTGTCTTATATATATGACAGATTTATGTATTAATCCAGAACTTGGACCTGCATCGTTAAATTT	455
Db	205	AGTGTCTTATATATGACAGATTTATGTATTAATCCAGAACTTGGACCTGCATCGTTAAATTT	146
OY	456	ATTATGTGTAACTTCAATATGTGTGCACTTAATATATCTCCACACATTAATCTGAAAAA	515
Db	145	ATTATGTGTAACTTCAATATGTGTGCACTTAATATATCTCCACACATTAATCTGAAAAA	86
OY	516	CTGATTTTGTATGTAAGAGCTCCCTTTCTATTTACTAGTCTGTGTACATACATCTTTT	575
Db	85	CTGATTTTGTATGTAAGAGCTCCCTTTCTATTTACTAGTCTGTGTACATACATCTTTT	26
OY	576	TATGAGCTATGAATAAACAATTTT 600	
Db	25	TATGAGCTATGAATAAACAATTTT 1	

```

RESULT 10
US-09-998-598-.235/c
; Sequence 235, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:

```

```

1  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
2  TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
3  FILE REFERENCE: 210121.561
4  CURRENT APPLICATION NUMBER: US/09/998,998
5  CURRENT FILING DATE: 2001-11-16
6  NUMBER OF SEQ ID NOS: 2606
7  SOFTWARE: Corixa Invention Disclosure Database
8  SEQ ID NO 235
9  LENGTH: 563
10 TYPE: DNA
11 ORGANISM: Homo sapiens
12 FEATURE:
13 NAME/KEY: misc_feature
14 LOCATION: 28, 109, 266, 332, 368, 506, 512, 531
15 OTHER INFORMATION: n = A,T,C or G
16 US-09-998-598-235

```

Query Match	65.4%	Score 439.4	DB 10	Length 563
Best Local Similarity	98.7%	Pred. No. 5	8e-66	
Matches 440	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	156	TCGATGAAGGAGATGGGTGACATGCATTCATGGGAACTCATGCATCTATATATTTGAA	215	
Db	451	TCGGATTAAGGAGCATGGGTGACATGCATTCATGGGAACTCAATGCATTAATATATTTTGA	392	

QY 216 GATTATAGAGAAGGAAATAGCAATGACACAAATTACAATGTGTGCGTGAGC 275
 |||||||
 Db 391 GATTATAGAGAAGGAAATAGCAATGACACAAATTACAATGTGTGCGTGAGC 332
 QY 276 AAGACATCTTTGAAGCTCAGTGTGTAGTTTACATCATATATTTGTATAGTGA 335
 |||||||
 Db 331 AAGACATCTTTGAAGCTCAGTGTGTAGTTTACATCATATATTTGTATAGTGA 272
 QY 336 CCTGACTCAAAATATTAAGCAGCTTGAACCTGCTTACCAATCTTGAATTTGACACA 395
 |||||||
 Db 271 CCTGCTGCAAAATATTAAGCAGCTTGAACCTGCTTACCAATCTTGAATTTGACACA 212
 QY 396 AGTGTCTTATATGACAGATCTAATGTAAATCCAGAACTGGATCCTCATGTTAAAT 455
 |||||||
 Db 211 AGTGTCTTATATGACAGATCTAATGTAAATCCAGAACTGGATCCTCATGTTAAAT 152
 QY 456 ATTTATGTGTATACATTCAAATGTGTGCAATTAATATGCTCCAGATTAATCTGA 515
 |||||||
 Db 151 ATTTATGTGTATACATTCAAATGTGTGCAATTAATATGCTCCAGATTAATCTGA 92
 QY 516 CTGATTTGTGATGAAGCTGCTTCTTATTTACTTACTGAGTCTGTACATACATCTTTT 575
 |||||||
 Db 91 CTGATTTGTGATGAAGCTGCTTCTTATTTACTTACTGAGTCTGTACATACATCTTTT 32
 QY 576 TATGAGCTATGAATTAACATTTTA 601
 |||||||
 Db 31 TATNAGCTATGAATTAACATTTTA 6

RESULT 11

US-09-777-564-1254/c
 ; Sequence 1254, Application US/09777564
 ; Patent No. US20020022591A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.493
 ; CURRENT APPLICATION NUMBER: US/09/777,564
 ; CURRENT FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 1730
 ; SOFTWARE: FastSeq for Window Version 4.0
 ; SEQ ID NO 1254
 ; LENGTH: 439
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-777-564-1254

Query Match 65.3%; Score 439; DB 10; Length 439;
 Best Local Similarity 100.0%; Pred. No. 6.6e-66;
 Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AAGGAGATGGGTGAGATGATAGGACATCGCATCTATATATTTGAAGATTAT 221
 |||||||
 Db 439 AAGGAGATGGGTGAGATGATAGGACATCGCATCTATATATTTGAAGATTAT 380
 QY 222 AGAAGAAGGAAATAGCAATGACACAAATTAACAATGTGTGCGTGAGGAGACA 281
 |||||||
 Db 379 AGAAGAAGGAAATAGCAATGACACAAATTAACAATGTGTGCGTGAGGAGACA 320
 QY 282 TCTTTGAGGTGATGAGTGTAGTTTAACATCATATATTTGTAACTGAACCTGTA 341
 |||||||
 Db 319 TCTTTGAGGTGATGAGTGTAGTTTAACATCATATATTTGTAACTGAACCTGTA 260
 QY 342 CTGAATAATATAGCAGCTTGAACCTGCTTACCATCTTGAATTTGACCAAGTGTG 401
 |||||||
 Db 259 CTGAATAATATAGCAGCTTGAACCTGCTTACCATCTTGAATTTGACCAAGTGTG 200
 QY 402 TTATATATGAGATCTAATGTAATAATCCAGAACTGGACTCCATCGTTAAATTTATTTAT 461
 |||||||
 Db 199 TTATATATGAGATCTAATGTAATAATCCAGAACTGGACTCCATCGTTAAATTTATTTAT 140

QY 462 GTGTACATTCAAATGTGTGATTAATATGCTTCCACAGTAATAATCTGAATAACTGATT 521
 |||||||
 Db 139 GTGTACATTCAAATGTGTGATTAATATGCTTCCACAGTAATAATCTGAATAACTGATT 80
 QY 522 TGTGATGAAGAGCTGCTTCTATTTACTTACTGAGTCTGTACATACATCTTTTATAG 581
 |||||||
 Db 79 TGTGATGAAGAGCTGCTTCTATTTACTTACTGAGTCTGTACATACATCTTTTATAG 20
 QY 582 CTATGAATTAACATTTT 600
 |||||||
 Db 19 CTATGAATTAACATTTT 1

RESULT 12

US-09-998-598-290/c
 ; Sequence 290, Application US/09998598
 ; Patent No. US20020150922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Chenault, Ruth A.
 ; APPLICANT: Meagher, Madeline Joy
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.561
 ; CURRENT APPLICATION NUMBER: US/09/998,598
 ; CURRENT FILING DATE: 2001-11-16
 ; NUMBER OF SEQ ID NOS: 2606
 ; SOFTWARE: Corixa Invention Disclosure Database
 ; SEQ ID NO 290
 ; LENGTH: 534
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1, 403
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-998-598-290

Query Match 64.5%; Score 433.4; DB 10; Length 534;
 Best Local Similarity 99.6%; Pred. No. 5.9e-65;
 Matches 445; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 156 TCAGATTAAGGAGATGGGTGAGATGATAGGAACTCAATGCAATTC-TATATATTTGA 214
 |||||||
 Db 450 TCAGATTAAGGAGATGGGTGAGATGATAGGAACTCAATGCAATTC-TATATATTTGA 391
 QY 215 AGATTATAGAGAAGGAAATAGCAATGACACAAATTAACAATGTGTGCGTGAGC 274
 |||||||
 Db 390 AGATTATAGAGAAGGAAATAGCAATGACACAAATTAACAATGTGTGCGTGAGC 331
 QY 275 GAAGACATCTTTAAGTCTAGTGTGTGTAGTTTAACATCATATATTTGTAACTGA 334
 |||||||
 Db 330 GAAGACATCTTTAAGTCTAGTGTGTGTAGTTTAACATCATATATTTGTAACTGA 271
 QY 335 ACCTGACTCAAAATATTAAGCAGCTTGAACCTGCTTTCACATCTTGAATTTGACCA 394
 |||||||
 Db 270 ACCTGACTCAAAATATTAAGCAGCTTGAACCTGCTTTCACATCTTGAATTTGACCA 211
 QY 395 AAGTGTCTTATATATGACATCTAATGTAAATCCAGAACTGGACTCCATCGTTAAAT 454
 |||||||
 Db 210 AAGTGTCTTATATATGACATCTAATGTAAATCCAGAACTGGACTCCATCGTTAAAT 151
 QY 455 TATTTATGTGTAACTCAATTAAGTGTGCAATTAATGTCTTCACAGTAAATCTGA 514
 |||||||
 Db 150 TATTTATGTGTAACTCAATTAAGTGTGCAATTAATGTCTTCACAGTAAATCTGA 91
 QY 515 ACTGATTTGTGATTTGAACCTGCTTCTATTTACTTACTGAGTCTGTATACATCTTTT 574
 |||||||
 Db 90 ACTGATTTGTGATTTGAACCTGCTTCTATTTACTTACTGAGTCTGTATACATCTTTT 31
 QY 575 TTATGAGCTATGAATTAACATTTTA 601

Db 30 TTATGAGCTATGAATAAATCATTTTA 4

RESULT 13

US-09-998-598-1423
 : Sequence 1423, Application US/09998598
 : Patent No. US20020150922A1
 : GENERAL INFORMATION:
 : APPLICANT: Stolk, John A.
 : APPLICANT: Xu, Jiangchun
 : APPLICANT: Chenault, Ruth A.
 : APPLICANT: Meagher, Madeleine Joy
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 : FILE REFERENCE: 210121.561
 : CURRENT APPLICATION NUMBER: US/09/998,598
 : NUMBER OF SEQ ID NOS: 2606
 : SOFTWARE: Corixa Invention Disclosure Database
 : SEQ ID NO 1423
 : LENGTH: 566
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 565
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-998-598-1423

Query Match Best Local Similarity 64.0%; Score 430.2; DB 10; Length 566;
 Matches 432; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 156 TCAGATAAAGAGATGGTGCATGATGCAATGCACTCAATGATATATTAATTGAA 215
 Db 131 TGAGATTAAGAGATGGTGCATGATGCACTCAATGATATATTAATTGAA 190
 QY 216 GATTATAGAGAAGGAATAGCAATGACACAAATTCGAATGCTGGTGGGAGC 275
 Db 191 GATTATAGAGAAGGAATAGCAATGACACAAATTCGAATGCTGGTGGGAGC 250
 QY 276 AAGACATCTTTGAAGTCAATGATGCTGATGCTTAAATCAATATATTTGTAATGAAA 335
 Db 251 AAGACATCTTTGAAGTCAATGATGCTGATGCTTAAATCAATATATTTGTAATGAAA 310
 QY 336 CCTGACTCAAAATATATAGCAGCTGAAACGCGCTTTACCAATCTTGAATTTGACCACA 395
 Db 311 CCTGACTCAAAATATATAGCAGCTGAAACGCGCTTTACCAATCTTGAATTTGACCACA 370
 QY 396 AATGCTTATATATGAGATCTAATGCTAATTAATCCAACTGAGCTCCATCGTTAAATT 455
 Db 371 AATGCTTATATATGAGATCTAATGCTAATTAATCCAACTGAGCTCCATCGTTAAATT 430
 QY 456 ATTATATGCTAATCAATGCTGATGCTTAAATATGCTTCCACAGTAATAATCTGAAAA 515
 Db 431 ATTATATGCTAATCAATGCTGATGCTTAAATATGCTTCCACAGTAATAATCTGAAAA 490
 QY 516 CTGATTTGATTTGAAGAGCTGCTTTCTATTACTGATGCTTTGATACATACACTTTT 575
 Db 491 CTGATTTGATTTGAAGAGCTGCTTTCTATTACTGATGCTTTGATACATACACTTTT 550
 QY 576 TATGAGCTATGAATA 591
 Db 551 TATGAGCTATGAATA 566

RESULT 14
 US-09-922-217-591/c
 : Sequence 591, Application US/09922217
 : Patent No. US200207641A1
 : GENERAL INFORMATION:
 : APPLICANT: Xu, Jiangchun

APPLICANT: Lodes, Michael J.
 : APPLICANT: Secretist, Heather
 : APPLICANT: Benson, Darin R.
 : APPLICANT: Meagher, Madeleine Joy
 : APPLICANT: Stolk, John A.
 : APPLICANT: Wang, Tonglong
 : APPLICANT: Jiang, Yugu
 : APPLICANT: Smith, Carole Lynn
 : APPLICANT: King, Gordon E.
 : APPLICANT: Wang, Aijun
 : APPLICANT: Clapper, Jonathan D.
 : TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 : FILE REFERENCE: 210121.471C13
 : CURRENT APPLICATION NUMBER: US/09/922,217
 : NUMBER OF SEQ ID NOS: 1124
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 591
 : LENGTH: 425
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-922-217-591

Query Match Best Local Similarity 61.8%; Score 415.4; DB 10; Length 425;
 Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 156 TCAGATAAAGAGATGGTGCATGATGCAATGCACTCAATGATATATTAATTGAA 215
 Db 417 TGAGATTAAGAGATGGTGCATGATGCACTCAATGATATATTAATTGAA 358
 QY 216 GATTATAGAGAAGGAATAGCAATGACACAAATTCGAATGCTGGTGGGAGC 275
 Db 357 GATTATAGAGAAGGAATAGCAATGACACAAATTCGAATGCTGGTGGGAGC 298
 QY 276 AAGACATCTTTGAAGTCAATGATGCTGATGCTTAAATCAATATATTTGTAATGAAA 335
 Db 297 AAGACATCTTTGAAGTCAATGATGCTGATGCTTAAATCAATATATTTGTAATGAAA 238
 QY 336 CCTGACTCAAAATATATAGCAGCTGAAACGCGCTTTACCAATCTTGAATTTGACCACA 395
 Db 237 CCTGACTCAAAATATATAGCAGCTGAAACGCGCTTTACCAATCTTGAATTTGACCACA 178
 QY 396 AATGCTTATATATGAGATCTAATGCTAATTAATCCAACTGAGCTCCATCGTTAAATT 455
 Db 177 AATGCTTATATATGAGATCTAATGCTAATTAATCCAACTGAGCTCCATCGTTAAATT 118
 QY 456 ATTATATGCTAATCAATGCTGATGCTTAAATATGCTTCCACAGTAATAATCTGAAAA 515
 Db 117 ATTATATGCTAATCAATGCTGATGCTTAAATATGCTTCCACAGTAATAATCTGAAAA 58
 QY 516 CTGATTTGATTTGAAGAGCTGCTTTCTATTACTGATGCTTTGATACATACACTTT 572
 Db 57 CTGATTTGATTTGAAGAGCTGCTTTCTATTACTGATGCTTTGATACATACACTTT 1

RESULT 15
 US-09-833-263-591/c
 : Sequence 591, Application US/09833263
 : Patent No. US20020110547A1
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Aijun
 : APPLICANT: Clapper, Jonathan D.
 : APPLICANT: Stolk, John A.
 : APPLICANT: Meagher, Madeleine J.
 : TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 : FILE REFERENCE: 210121.471C12
 : CURRENT APPLICATION NUMBER: US/09/833,263
 : NUMBER OF SEQ ID NOS: 1093
 : SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 591
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-833-263-591

Query Match 61.8%; Score 415.4; DB 10; Length 425;
Best Local Similarity 99.8%; Pred. No. 6e-62;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	156	TCGATTAAGAGATGGGTGAGATGCAATAGGAACTCAATGCACTAATATTAATTGAA	215
DB	417	TCGATTAAGAGATGGGTGAGATGCAATAGGAACTCAATGCACTAATATTAATTGAA	358
QY	216	GATTATAGAGAAGGAAATAGCAAAATGACACAAATTACAAATGTGTGCGTGGAGC	275
DB	357	GATTATAGAGAAGGAAATAGCAAAATGACACAAATTACAAATGTGTGCGTGGAGC	298
QY	276	AACACATCTTTGAAGCTCATGAGTTGTTAGTTAACATCATATATTGTAATAGTGAA	335
DB	297	AACACATCTTTGAAGCTCATGAGTTGTTAGTTAACATCATATATTGTAATAGTGAA	238
QY	336	CCTGTACTCAAAATATTAAGGAGCTTGAAGCTTACCAATCTGAAATTTGACCACA	395
DB	237	CCTGTACTCAAAATATTAAGGAGCTTGAAGCTTACCAATCTGAAATTTGACCACA	178
QY	396	AGTGTCTTAATATATGAGATCTAATGTAAATCCAGAACTTGGACTCCATCGTTAAATT	455
DB	177	AGTGTCTTAATATATGAGATCTAATGTAAATCCAGAACTTGGACTCCATCGTTAAATT	118
QY	456	ATTATGTGTAAACATCAATGTGTGATTAATATGCTTCCACAGTAATAATCTGAAAA	515
DB	117	ATTATGTGTAAACATCAATGTGTGATTAATATGCTTCCACAGTAATAATCTGAAAA	58
QY	516	CTGATTTGTATGAAAGCTGCTTCTATTACTTGAGTCTGTACATACATCTT	572
DB	57	CTGATTTGTATGAAAGCTGCTTCTATTACTTGAGTCTGTACATACATCTT	1

Search completed: November 26, 2002, 23:47:34
Job time : 28.289 secs